Jonna Kela

Eurooppalainen tekoäly

Millaista oikeudellista alustaa Euroopan unioni luo tekoälylle?

Johtamisen yksikkö
Julkisoikeus, Pro gradu -tutkielma
Hallintotieteiden maisteri

Vaasa 2023

VAASAN YLIOPISTO
Johtamisen akateeminen yksikkö
Tekijä: Jonna Kela
Tutkielman nimi: Eurooppalainen tekoäly: Millaista oikeudellista alustaa Euroopan unioni luo tekoälylle?
Tutkinto: Hallintotieteiden maisteri
Oppiaine: Julkisoikeus
Työn ohjaaja: Kristian Siikavirta
Valmistumisvuosi: 2023 **Sivumäärä: 71**

TIIVISTELMÄ:
Tekoäly on vasta viimeisen vuosikymmenen aikana näkyvästi yleistynyt teknologia eikä sen varalle olla vielä nimenomaista lainsäädäntöä annettu. Tekoälyn nähdään kuitenkin aiheuttavan merkittäviä ja kriittisiä riskejä yksilön elämään ja yhteiskuntaan, minkä takia Euroopan unioni on ryhtynyt jo vuonna 2018 toimiin tekoälyä koskevien perus- ja ihmisoikeuksia kunnioittavien puitteiden luomiseksi Euroopassa. Eräs tärkeä askel tämän tavoitteen toteutumiseksi on tekoälyn määrittely oikeudellisessa mielessä. Vuonna 2022 antamansa tekoälysäädösehdotuksen turvin unionin aikomus on antaa lopullinen oikeudellinen olemus tekoälylle. Unionin lähteisiin perustuen voidaan tekoälyn sanoa olevan autonominen kokonaisuus, joka koostuu algoritmeista ja datasta. Tekoälyn määritelmä on tarkoitettu jätettävän mahdollisimman teknologianeutraaliksi, jotta tuleva lainsäädäntö, ohjeet ja suositukset vastaisivat paremmin teknologisen kehityksen tahtia. Säädösehdotusten lisäksi EU on tehnyt mittavaa työtä julkaistakseen erilaisia suosituksia, ohjeita ja muuta materiaalia luotettavan ja eettisen tekoälyn kehittämisen tueksi.

AVAINSANAT: tekoäly, data, Euroopan unioni, koneoppiminen, ihmiskeskeisyys, data-avaruus, etiikka

Sisällys

1 Johdanto 5

1.1 Tutkielman taustaa 5

1.2 Tutkimuksen kohde ja menetelmä 9

2 Mitä on tekoäly? 12

2.1 Nykyaikainen tekoäly oppii itsenäisesti 13

2.2 Tekoäly lain silmin 18

2.2.1 Tekoälyn etiikka 19

2.2.2 Luotettavan tekoälyn arviointi (ALTAI) 22

2.2.3 Ihmiskeskeisyys tekoälyn periaatteena 29

2.2.4 Suuririskiset tekoälyjärjestelmät 31

3 Tekoälyn vastuuasettelu 36

3.1 Vastuu tekoälyn tekijällä 36

3.2 Vahingonkorvausvastuu 40

3.3 Käyttäjän ja tarjoajan kantama vastuu 42

3.4 Euroopan unionin rekisteri tekoälyjärjestelmille 45

4 Tekoäly ja Data 46

4.1 GDPR ja kyberturvallisuusasetus 47

4.2 Tiedonhallintalaki 49

4.3 Millaista dataa voidaan käyttää tekoälyn muovaamiseen? 50

4.4 Euroopan unionin yhteinen data-avaruus 52

4.5 Tietojen käsittelyn periaatteet 56

5 Tietosuoja perusoikeutena 59

5.1 Perusoikeuskirja ja perustuslaki 60

5.2 Henkilötietojen suoja ja tietosuoja 62

5.3 Tietoturva ja kyberturva 63

6 Johtopäätökset 65

Lähteet 68

Kuvat

Kuva 1. kasvojentunnistusjärjestelmä DNN-tekniikalla.

Kuviot

Kuvio 1. Euroopan unionin julkaisuja eettiselle tekoälylle.

Lyhenteet

AGI=Yleinen tekoäly (Artificial general intelligence)
AI=Tekoäly (Artificial intelligence)
AI HLEG=Tekoälyä käsittelevä korkean tason asiantuntijaryhmä
ALTAI=Luotettavaa tekoälyä koskeva arviointilista
ANN=Neuroverkko (Artificial neural network)
DNN=Syvä neuroverkko (Deep neural network)
DPIA=Tietosuojaa koskeva vaikutustenarviointi
EU=Euroopan unioni
GDPR=Yleinen tietosuoja-asetus
HITL=human in the loop
HOC=human on command
HOTL=human on the loop
MLP=monikerroksinen perseptroniverkko (Multilayer perceptron)
SEU=Sopimus Euroopan unionista
SEUT=Sopimus Euroopan unionin toiminnasta
TAR=taktinen lisätty todellisuus (Tactical augmented reality)
USACM=The US Public Policy Council od the Association for Computing Machinery
YK=Yhdistyneet kansakunnat

1 Johdanto

1.1 Tutkielman taustaa

Kokeellisessa psykologiassa älykkyydellä viitataan yleisiin kognitiivisiin valmiuksiin, joita mitataan standardoiduilla päättelytehtävillä. Kognitio taasen tarkoittaa niitä monimutkaisia prosesseja ja toimintoja ja niiden rakenteita, joihin älykäs käyttäytyminen perustuu. Kognition tutkimus siis pyrkii ymmärtämään, millaiset prosessit mahdollistavat älykkään toiminnan. Osasyy tekoälyn kehitykselle onkin ollut yritys ymmärtää ihmisen ajattelua ja näin ollen myös kognitiotiede syntyi laskennallisen näkökulman soveltamisesta ihmisälyn tutkimiseen.[1]

Tekoälyn (AI) yhteydessä älykkyydellä viitataan usein älykkääseen käyttäytymiseen, joka ilmenee joustavuudella ja tarkoituksenmukaisella toiminnalla monimutkaisessa, muuttuvassa ja ennustamattomassa ympäristössä. Määritelmä korostaa älykkään toimijan ja ympäristön keskinäistä suhdetta älyn määrittelemiseksi. Älykkyys siis ilmenee tämän mukaan siinä, miten toimija kykenee sopeutumaan tiedon käsittelyn varaista toimintaansa monimutkaisessa ympäristössä.[2] Tekoälyn olennaisimpia ominaisuuksia ovat oppivuus, suorituskyvyn laajamittaisuus sekä autonomisuus. Tekoäly ei siis ole ohjelmisto, joka on ennalta ohjelmoitu suorittamaan tietty tehtävä, vaan se oppii itsenäisesti suoriutumaan useista tehtävistä ympäristönsä muutosten alla. Mitä laajempi tehtävä tekoälylle halutaan antaa, sitä enemmän vaaditaan oppivuutta. Autonomia määrittelee, kuinka paljon tekoälyä täytyy opettaa etukäteen ratkaisemaan määritelty ongelma ja kuinka hyvin tekoäly kykenee itse määrittelemään ratkaistavan ongelman ja tuottamaan siihen soveltuvan ratkaisun.[3] Tekoälyä vielä edistyneempää keinotekoista älykkyyttä on yleinen tekoäly (AGI), jonka tarkoitus on simuloida ihmisaivojen kaltaista tietoisuutta.

[1] Lappi ja muut, 2018, s.42-43.
[2] Lappi ja muut, 2018, s.43.
[3] Työ- ja elinkeinoministeriö, 2017, s.62-63.

Tällainen tekoäly on vahvaa tekoälyä, kun taas moderni tekoäly kuvataan heikoksi tekoälyksi. [4]

Vaikka tekoälyä on tutkittu ja sovellettu 1950-luvulta lähtien, vasta viime aikoina sitä on alettu kehittämään ja hyödyntämään tosiasiallisesti. Syynä on yksinkertaisesti laskentakapasiteetin ja datan saatavuus, joita kumpaakaan ei ole vielä edellisen vuosisadan aikana ollut saatavilla tarpeellisissa määrin.[5] Vuonna 2018 Nvidia -yhtiö uutisoi syväoppimisjärjestelmästä, jonka avulla robotti alkoi koeoloissa jäljittelemään ihmisen tekemistä vain tarkkailemalla ja matkimalla. Kyseinen robotti hyödynsi tekniikkanaan konenäköä ja hahmontunnistusta, joiden avulla mahdollistui siirtää oikeanväriset palikat oikeassa järjestyksessä torniksi esimerkin mukaisesti. Järjestelmä ei ollut tässä tapauksessa riippuvainen ympäristöstä tai pohjatiedoista. Toisena esimerkkinä Japanin kansallisen informatiikkainstituutin professori Noriko Arai perusti vuonna 2011 projektin, jonka tarkoituksena oli saada tekoälyrobotti niin ajattelevaksi ja tietäväksi, että se kykenisi läpäisemään japanin yliopistojen vakioidut harjoituskokeet ja Tokion yliopiston oikeat pääsykokeet. Robotin ongelmaksi kuitenkin koitui tehtävät, jotka vaativat päättelyä, luovuutta sekä asioiden merkityksen ja tarkoituksen kriittistä ymmärtämistä.[6]

Vaikka tekoäly ei tähän päivään mennessä sovellu ihmisälyn korvaajaksi, pystytään tätä teknologiaa hyödyntämään useissa eri paikoissa avustamaan ihmistä. Esimerkiksi New Yorkin syöpäsairaalassa ”Memorial Sloan-Kettering Cancer Center” on käytetty IBM watson-tekoälyä diagnoosien laatimiseen ja hoitoprosessien suunnitteluun. Watsonille on tässä tapauksessa syötetty potilastietoja, joiden perusteella tekoäly on verrannut kyseisiä tietoja jo aiemmin annettuihin historiaraportteihin, lääketieteellisiin julkaisuihin sekä potilasrekisteritietoihin. Tekoälyä on hyödynnetty USA:ssa myös juridisiin toimenpiteisiin ja istuntojen valmisteluun. Tässä tekoäly analysoi juridisia dokumentteja ja tuotti niiden perusteella aineistoa valmisteluprosessiin.[7] Myös vuonna 2022 julkaistua tekoälybottia

[4] Siukonen ja Neittaanmäki, 2019, s.29 ja s.43.
[5] Työ- ja elinkeinoministeriö, 2017, s.15.
[6] Siukonen ja Neittaanmäki, 2019, s.30-31.
[7] Virtanen ja Stenvall, 2014, s.106.

ChatGPT:tä on hyödynnetty juridiseen valmistelutyöhön. Johtuen chatbotin luonteesta, tekoäly oli kuitenkin luonut ja esittänyt lakimiehelle viittauksia täysin keksittyihin oikeudellisiin tapauksiin.[8]

Automaatiolla, kuten tekoälyllä, pyritään vähentämään manuaalista työtä ja tätä kautta inhimillisiä virheitä sekä vapauttamaan ihmisen käyttämää aikaa enemmän arvoa tuottaviin tehtäviin ja täten lisäämään tuottavuutta ja tyytyväisyyttä. Yleisesti ottaen tekoäly parhaimmassa tapauksessa mahdollistaa ihmisen apuvälineenä tehtävien laadukkaamman ja nopeamman suorittamisen.[9] Esimerkiksi Valtion talous- ja henkilöstöhallinnon palvelukeskus ”Palkeet” hyödyntää tekoälyä ostolaskujen tiliöintiin. Palkeet tuottaa tiliöintipalveluja noin 50:lle valtion organisaatiolle, mikä luonnollisesti aiheuttaa suuria työmääriä kyseiselle virastolle. Tekoälyä käyttämällä pyrkimys on suoraan vähentää tiliöintiin liittyvää manuaalityötä niin, että tekoäly täydentää verkkolaskulle tiliöintiennusteen, jonka jälkeen lasku palautuu normaaliin käsittelyprosessiin.[10] Myös kansalaisille voidaan tekoälyn avulla mahdollistaa digitaalisia palveluja sujuvammin ja oikeaan aikaan[11].

Tekoälyn tarjoamista mahdollisuuksista huolimatta, tutkijat varoittavat sen riskeistä erityisesti liittyen ”musta laatikko” -luonteenomaisuuteen. Tekoälyjärjestelmiin saattaa joutua vaikeasti löydettäviä virheitä ohjelmien monimutkaisuuden, ohjelmointivirheiden, perusoletusten ja erilaisten tekniikoiden kautta[12]. Tekoälyn kanssa esimerkiksi päätöksenteko ei välttämättä ole vaaditussa määrin läpinäkyvää, järjestelmä saattaa alkaa soveltamaan muun muassa sukupuolisia vinoumia ja järjestelmät voivat itsessään olla tai johtaa yksityiselämän loukkaukseen[13]. Näin ollen suositus on, että tällaisia mustia laatikoita mahdollisesti sisältäviä järjestelmiä ei käytettäisi tai käytettäisiin erityistä huolellisuutta noudattaen lakien toimeenpanossa, terveydenhuollossa ja koulutuksessa.

[8] BBC, 2023.
[9] Kaarlejärvi ja Salminen, 2018, s.182-183 ja s.22-23.
[10] Palkeet, 2021.
[11] Työ- ja elinkeinoministeriö, 2017, s.52.
[12] Siukonen ja Neittaanmäki, 2019, s.55-56.
[13] Euroopan komissio, 2020, s.1.

Tekoälyn käyttöä ja käyttöönottoa koskevat ohjeet eivät tällä hetkellä ole pakottavia eikä lainsäädäntöäkään ole varustettu tekoälyn sääntelemiseksi.

Mikäli tekoäly aiheuttaisi virheitä tai rikoksia, lainsäädännöllisesti sovellettaisiin muun muassa vahingonkorvauslakia, kuluttajansuojalakia ja kauppalakia sekä sopimuksiin kirjattuja ehtoja. Lisäksi tekoälyä varten on tehnyt suosituksia muun muassa Yhdysvaltalainen AI now Institute -tutkimuslaitos. Suositusten mukaan tekoälyalgoritmeista tulisi muun muassa tehdä puolueettomia ja mahdollisimman neutraaleja ja ennen käyttöönottoa tulisi suorittaa tarkkoja testejä, jotka varmistavat puolueettomuuden. Lisäksi tekoälyjärjestelmiä tulisi voida julkisesta seurata ja arvioida ja kehityksestä tulisi hyödyntää monitieteisyyttä, poikkitieteellisyyttä, standardointia ja muidenkin kuin teknisen alan näkökulmia. Toisena esimerkkinä the US Public Policy Council of the Association for Computing Machinery eli USACM on julkaissut ohjelman läpinäkyvyydelle ja vastuullisuudelle.[14]

Myös Euroopan komissio on ottanut työkseen varmistua, että uusi sovellettava teknologia tehdään palvelemaan eurooppalaisia ja parantamaan unionin kansalaisten elämää kunnioittaen kunkin oikeuksia ja velvollisuuksia[15]. Eurooppalainen lähestymistapa tekoälylle on ihmiskeskeisyys ja luotettavuus niin, että tekoälyä kehitetään luotujen sääntöjen pohjalta markkinoiden, julkisen sektorin ja ihmisten turvallisuuden sekä perusoikeuksien toteutumista varmistaen. Vuodesta 2018 lähtien unioni on vienyt eteenpäin ideaa turvallisen tekoälyn käyttöönotolle Euroopassa. Eräitä olennaisia julkaisuja ovat: ”Tekoäly Euroopalle”, ”Tekoälyä koskeva koordinoitu suunnitelma”, ”Eettiset ohjeet luotettavalle tekoälylle”, ”Valkoinen kirja tekoälystä”, ”Luotettavaa tekoälyä (ALTAI) koskeva arviointiluettelo” sekä ”Ehdotus asetukseksi tekoälyä koskevien yhdenmukaistettujen sääntöjen vahvistamiseksi”. [16]

[14] Siukonen ja Neittaanmäki, 2019, s.55-56.
[15] Euroopan komissio, 2020, s.2
[16] Euroopan komissio, A European approach to artificial intelligence.

1.2 Tutkimuksen kohde ja menetelmä

Tutkielmani tarkoitus on selventää ja koota yhteen tekoälyn kehittämiseen vaikuttavia seikkoja oikeudellisesta näkökulmasta sekä, mitkä seikat luovat rajoitteita tekoälyn kehityksessä. Tutkimuskysymykseni ovat seuraavat:

1. Mitä tekoäly on?
2. Mikä luo rajoitteita tekoälylle ja miten kyseiset rajoitteet näkyvät tekoälyjärjestelmien kehityksessä?

Aloitan tutkielmani esittämällä tekoälylle määritelmän niin teknisessä kuin oikeudellisessa mielessä samalla valottaen, miten Euroopan unionin asettamat ehdot luovat tekoälylle rajoitteita ja velvoitteita lain silmin. Unionin ehdot luovat tavoitteen tekoälyn olemukselle ja toiminnallisuudelle esimerkiksi asettamalla tiettyjä eettisiä velvoitteita. Lisäksi tekoälyn määrittämiseksi voidaan hyödyntää muun muassa tekoälyn kyvykkyyksien, käyttökohteiden ja riskien ulottuvuutta. Näin ollen haluan avata, mistä on kyse, kun puhutaan tekoälyä koskevista eettisistä ohjeista sekä, mitä tarkoitetaan järjestelmän suuririskisyyden arvioinnilla.

Euroopan unionin tekoälylle antamat vaatimukset asettavat tietyille toimijoille vastuuta tekoälyn eri elinkaaren ajoilla. Samoin olemassa oleva ja tuleva lainsäädäntö tulee myös asettamaan näitä vaatimuksia, joita tekoälyn kanssa tekemisissä olevien on huomioitava. Koska tekoälyyn liittyvät vaatimukset ovat vahvasti yhteydessä vastuukysymyksiin virheiden sattuessa, haluan myös lyhyesti avata tätä aiheitta ja sitä, mitä kunkin toimijan vastuualueeseen kuuluu siitäkin huolimatta, ettei kyseessä olisi oikeudellisesti määriteltävä vastuuvelvoite.

Lopuksi otan esille datan merkityksellisyyden ja siihen liittyviä rajoitteita ja velvoitteita. Omina kappaleina kerron, millä tavalla yleinen tietosuoja-asetus ja kyberturvallisuusasetus vaikuttavat henkilötietojen käyttöön tekoälyn opetusdatana; miten tiedonhallintalaki ottaa kantaa henkilötietojen käsittelyyn; sekä yleisesti, millaista dataa tekoälyn opetuksessa tulisi hyödyntää ja miten Euroopan unioni pyrkii edistämään datan saatavuutta.

Tutkimusmenetelmänä käytän oikeusdogmatiikkaa eli lainoppia, jonka tarkoituksena on tuottaa systemaattista, objektiivista ja perusteltua tietoa lain soveltamisen tarpeisiin. Tämä tapahtuu oikeussäännösten systematisoinnin ja tulkinnan avulla. Tulkinnan avulla luodaan merkityssisältöä lakiteksteille ja muulle oikeudelliselle tekstille. Systematisointi sen sijaan viittaa oikeussäännösten järjestämistä ristiriidattomaksi järjestelmäksi niin, että säännökset samalla asetetaan tiettyyn tulkintakontekstiin. Toteutetun tutkimuksen tavoite on siis tuottaa oikeudesta järjestynyt ja ristiriidaton normikokonaisuus. Oikeusnormilla tarkoitetaan oikeussäännöksen merkityssisältöä, joka syntyy säännöksen kielellisen ilmiasun tulkinnan seurauksena. Tulkintaa tehdään yleiskielellisesti sekä etsimällä kontekstin merkityksiä, mikä onnistuu sääntelykohteen määrittämisellä. Tulkinnassa on huomioitava myös muussa lainsäädännössä omaksuttu terminologia.

Eräs tulkinnan peruste on oikeuslähteen antama informaatio, jolloin argumenttina voidaan käyttää normiauktoriteetin todellista historiallista tarkoitusta tai säännöksen objektiivista tarkoitusta tulkittuna koko oikeusjärjestyksen tai sen määrätyn osan tavoitteiden kannalta. Tällaisen juridisen tulkinnan yhteydessä voidaan käyttää analogia -argumenttia. Analogialla tarkoitetaan, että arvioinnin kohteena oleva tilanne on niin samantyyppinen kuin jokin toinen tilanne, että näitä tilanteita arvioidaan samalla tavalla. Lisäksi tulkinta voi perustua arvoihin ja tavoitteisiin, jotka ovat johdettavissa yhteiskunnasta tai oikeussäännöksestä itsestään.[17] Erityisesti tekoälyn tilanteessa on tälle kyseiselle tekniikalle asetettu selkeä arvo- ja tavoitepohja, joiden tästä syystä tulisi näkyä erilaisissa lain tulkinnan tilanteissa.

Oikeuslähteet voidaan jakaa vahvasti velvoittaviin, heikosti velvoittaviin ja sallittuihin oikeuslähteisiin. Vahva velvoittavuus tarkoittaa, että julkisen vallan käytössä kyseistä normia on noudatettava virkavastuun uhalla.[18] Näin ollen viranomaisten ja tuomioistuinten on aina perustettava ratkaisunsa kirjoitettuun lainsäädäntöön[19]. Vahvasti velvoittavia

[17]Andström, 2003, ”Tutkimuksen eri tasojen teoreettinen aines”.
[18]Andström, 2003, ”Tutkimuksen eri tasojen teoreettinen aines”.
[19] Määttä ja muut, 2012, s.9.

oikeuslähteitä ovat kotimaiset säädökset, perustuslaki ja tavallinen laki, joiden nojalla voidaan edelleen antaa asetuksia ja näitä alemman tasoisia säädöksiä sekä tavanomainen oikeus eli maantapa, jolla tarkoitetaan oikeudenalan vakiintunutta käytäntöä. Myös Euroopan unionin antamat eurooppaoikeudelliset normit ovat vahvasti velvoittavia oikeuslähteitä, ja osalla näistä on myös etusija Suomen kansalliseen säädäntöön nähden. Heikosti velvoittavien oikeuslähteiden sivuuttamisesta ei seuraa sanktiota. Tällaista lähteistöä on lainvalmisteluaineisto, joka kuvastaa lainsäätäjän tarkoitusta sekä ennakkoratkaisut, jotka luovat ennakoitavuutta ja yhdenvertaisuutta tulkinnalle. Sallittuja oikeuslähteitä käytetään, kun velvoittavista oikeuslähteistä saatu tieto ei auta pääsemään tulkintaratkaisuun. Näitä lähteitä ovat lainopillinen kirjallisuus, oikeusjärjestelmäämme verrattuna samankaltaiset oikeusjärjestykset sekä historiallinen oikeus.[20] Tutkielmassani vahvan aseman saa eurooppaoikeudelliset oikeuslähteet sikäli, kun unioni on ottanut tehtäväkseen luoda Euroopan unionista etevän tekoälyn keskiön. Näin ollen unioni pyrkii luomaan yhtenäiset oikeudelliset lähtökohdat kullekin jäsenvaltiolle tekoälyn käyttöönoton rohkaisemiseksi samoin kuin tekoälyn kaupallisen ja ei-kaupallisen leviämisen mahdollistamiseksi samanlaisin velvoittein ja oikeuksin. Euroopan unioni on myös luonut periaatteita, arvoja ja ohjeita siihen, millaista tekoälyn tulisi olla ja mihin tulevalla tekoälyä koskevalla säädännöllä pyritään.

[20]Andström, 2003, ”Tutkimuksen eri tasojen teoreettinen aines” ja Määttä ja muut, 2012, s.8-9.

2 Mitä on tekoäly?

Vuonna 2017 Työ- ja elinkeinoministeriön julkaisussa ”Suomen tekoälyaika” todettiin ettei tekoälylle ole täsmällistä määritelmää. Kyseisessä julkaisussa tekoälyllä kuitenkin viitattiin laitteeseen, ohjelmistoon tai järjestelmään, joka kykenee oppimaan ja tekemään päätöksiä samalla tavalla kuin ihminen[21].

Termi ”Artificial intelligence” (AI) on ensimmäisen kerran tullut esille John McCarthyn esittämänä ja tälle termille määritelmän antoi tiivistetysti Dave Gershgorn[22]. Gershgornin mukaan tekoäly on ohjelmisto tai tietokoneohjelma, jossa on oppimismekanismi ja se käyttää oppimaansa tietoa päätöksen tekemiseksi uudessa tilanteessa ihmisen tavoin. Tämän esitellyn määritelmän mukaisesti, jotta ohjelmisto voidaan nimetä tekoälyksi, tulee sen olla oppiva ja ennen kaikkea tehdä päätöksiä tälle koneelle täysin tuntemattomien muuttujien vallitessa ilman, että ohjelmisto on jo opetettu kyseiseen päätökseen tai se on ohjelmoitu suorittamaan kyseessä oleva toimeksianto.

Edellä kuvattu määritelmä kohtaa ongelman erityisesti ihmismäisen käyttäytymisen kohdalla; mikä käytös ja ajattelutapa voidaan luokitella ihmismäiseksi? [23] Kysymykseen ei ole yksinkertaisesti selkeää vastausta ottaen huomioon, ettei ihmismielen toimintamekanismeja olla pystytty selvittämään täydellisesti eikä täten myöskään pystytä toisintamaan ihmisaivoissa tapahtuvia prosesseja. Puutteesta huolimatta vuonna 1959 Alan Turing kehitti testin todistamaan, pystyykö kone ajattelemaan kuten ihminen[24]. Testin perusideana mukaan osallistuu kolme ihmistä, joista yksi on haastattelija ja kaksi muuta vastaavat haastattelijalle. Toinen haastateltavista on nainen ja toinen mies ja näistä toinen myös valehtelee omasta sukupuolestaan ja toinen kertoo totuuden. Haastattelijan on pystyttävä selvittämään kummankin sukupuoli[25]. Kaikki kolme osanottajaa sijaitsevat eri

[21] Työ- ja elinkeinoministeriö, 2017, s.15.
[22] Siukonen ja Neittaanmäki, 2019, s.25-28.
[23] De Bruyne ja Vanleenhove, 2021, s.2-4.
[24] De Bruyne ja Vanleenhove, 2021, s.2-4.
[25] De Bruyne ja Vanleenhove, 2021, s.2-4.

huoneissa näkemättä toisiaan apunaan näppäimistö ja näyttö. Turingin testin tapauksessa toisen vastaajan tilalle vaihdetaan tietokone. Mikäli kone pystyy huijaamaan ihmistä yhtä usein kuin alun perin ihminen on kyennyt samassa asemassa, voidaan todeta koneen ajattelevan kuten ihminen.[26]

Turingin testin heikkous on, että tässä tapauksessa koneen kykenevyys imitoimaan ihmisen älyä ei todista älykkyyttä. Vaikkei testiä sellaisenaan kritisoidaan, on sen kautta pystytty soveltamaan samankaltaisia testejä muun muassa chatboteille. Chatbotin tapauksessa myöskin tavoitteena on saada chatin toinen osapuoli luulemaan bottia ihmiseksi.[27] Nykyaikaisen tekoälyn sanotaan jäävän kauaksi siitä, mitä yleinen tai laaja älykkyys ihmisellä on. Sen sijaan tuo koneen tekoäly on vain ohjelmisto, joka on rakennettu vaikuttamaan siltä, että se pitäisi sisällään käyttökelpoista älykkyyttä.[28]

2.1 Nykyaikainen tekoäly oppii itsenäisesti

Tämänhetkisellä tekoälyllä viitataan keinotekoiseen kokonaisuuteen, jossa on kyvykkyys tehdä älykkäiksi miellettyjä tai älykkäiltä vaikuttavia toimintoja, joissa ei tarvitse ajatella käsitteitä tai olioiden merkityksiä. Tällainen toiminto on esimerkiksi puheentunnistus, jonka avulla kone voi parsia yhteen jopa rikkinäisiä ja puutteellisia sanoja tai lauseita ja muuntamaan ne selkokielelle ja täten esimerkiksi kirjoittaa lääkärin tai juristin saneleman kertomuksen.[29] Samoin puheentunnistusta voidaan käyttää antamaan koneelle ihmisen puhumia käskyjä ja hyödyntää tätä ominaisuutta esimerkiksi puhelinvastaajassa.

Erityisen ratkaiseva askel tekoälylle on ollut saada ohjelmisto oppimaan samoin periaattein kuin ihminen oppii eli yrityksen ja erehdyksen kautta sekä informaatiota tarjoamalla[30]. Toki järjestelmän oppimiskyky ei tietenkään ole täysin ihmiseen verrattava. Kone

[26] De Bruyne ja Vanleenhove, 2021, s.2-4.
[27] De Bruyne ja Vanleenhove, 2021, s.2-4.
[28] Siukonen ja Neittaanmäki, 2019, s.28.
[29] Alho ja muut, 2018, s.7 ja Siukonen ja Neittaanmäki, 2019, s.83-85.
[30] Siukonen ja Neittaanmäki, 2019, s.30-31.

vaatii nimenomaan suuria määriä dataa toisin kuin ihmislapsi vaatii oppiakseen selkeästi paljon vähemmän tietoa, koska ihminen on kykenevä järkeilemään ja sopeutumaan paremmin uutuuteen ilman ennakkoon annettua informaatiota. Tekoäly ei myöskään vielä tänä päivänä kykene ymmärtämään ja käsittämään kontekstia vaikkakin sillä on kyvykkyyttä tunnistaa ja vieläpä hyvinkin pieniä yksityiskohtia, jotka jäisivät helposti ihmiseltä huomaamatta.[31]

Tekoäly oppii itsenäisesti, mutta se voi oppia vain, mitä ihminen antaa sen oppia, toisin sanoen riippuen täysin, millaiset algoritmit koneelle on kirjoitettu[32]. Algoritmilla tarkoitetaan tässä yhteydessä matemaattista kuvausta, mitä järjestelmän tulee tehdä jonkin ongelman ratkaisemiseksi tai tehtävän hoitamiseksi. Esimerkiksi koodatessa tietokoneohjelmaa python-ohjelmointikielellä, kyseessä on algoritmi. Tällaisessa tilanteessa kone suoriutuu täsmälleen ohjelmoitujen mallien mukaisesti välittämättä toimeksiannon mielekkyydestä tai ristiriitatilanteista, eikä se missään välissä unohda alkuperäistä toimeksiantoaan.[33]

Sikäli, kun puhutaan nykyaikaisesta tekoälystä, jolla ei ole yleistä ihmisälyn kaltaista kyvykkyyttä ajatella vaan älykkäitä toimintoja, on kyseessä heikko tekoäly ja yleensä tarkemmin tällä tarkoitetaan koneoppimiseen perustuvaa ohjelmistoa. Heikko tekoäly siis mahdollistaa ohjelmistoa muuttamaan epäonnistuneita käytöstapoja eli järjestelmä oppii koneoppimismenetelmiä hyödyntäen[34]. Heikolla tekoälyllä varustettuina järjestelminä voidaan pitää esimerkiksi hakukoneita tai robotti-imuria ohjaavaa järjestelmää[35].

Koneoppiminen syntyi jo 1980-luvulla, kun ohjelmistot saatiin ennalta ohjelmoitujen sääntöjen sijaan päätymään haluttuun lopputulokseen yleistämällä ja ennustamalla annetusta datasta[36]. Koneoppimisessa ohjelmisto improvisoi algoritmilla ja kohdistaa

[31] Topol, 2019, s.92.
[32] Siukonen ja Neittaanmäki, 2019, s.30-31.
[33] Siukonen ja Neittaanmäki, 2019, s.43 ja 46.
[34] Alho ja muut, 2018, s.7 ja Siukonen ja Neittaanmäki, 2019, s.43.
[35] Siukonen ja Neittaanmäki, 2019, s.43.
[36] Lappi ja muut, 2018, s.44.

matemaattisia laskentakaavoja dataa vasten. Näin ollen ohjelma kykenee luomaan algoritmiin pohjautuen uusia toiminnallisuuksia ja käyttäytymismalleja ilman ennalta määriteltyjä sääntöjä kyseisiin toimintoihin. Koneoppimisen eri menetelmiin luetellaan ohjattu (supervised learning) ja ohjaamaton (unsupervised learning) oppiminen sekä vahvistusoppiminen (Reinforcement learning). Ohjatussa oppimisessa algoritmi oppii, minkä tuotoksen (output) sen tulee yhdistää syötteeseen (input)[37]. Esimerkiksi algoritmille kerrotaan suoraan papukaijan, koiran ja kissan kuvista, mitkä ovat koiria ja täten algoritmi oppii tunnistamaan koiran[38]. Ohjaamattomassa oppimisessa algoritmille annetaan vain syöte ilman tuotosta ja algoritmin on itse tunnistettava kaavamaisia malleja ja rakenteita datasta[39]. Tässä algoritmille ei siis aseteta oikeaa vastausta vaan algoritmi löytää ennalta määrittelemättömän vastauksen itse. Vahvistusoppimisen kautta ”älykäs agentti” oppii optimaalisen joukon toimintoja, jotta se saavuttaa tietyn tavoitteen. Algoritmille ei kerrota, miten tuo tehdään, mutta se vastaanottaa suorituksistaan palautetta ohjaamaan oppimista[40]. Voidaan esimerkiksi asettaa algoritmille tavoitteeksi läpäistä pelissä tietty taso ja täten algoritmi kokeilemalla alkaa muodostaa optimaalisinta tapaa tähän[41].

Eräs tapa luoda koneälyä on jo 1940- luvulla syntyneet neuroverkot (Artificial neural network eli ANN). Neuroverkkojen idea on osittain jäljitellä luonnollisia ihmisen hermoverkkoja ja neuroneita keinotekoiseksi kokonaisuudeksi, jossa tietty syöte luo tietyn tuotoksen.[42] Neuroverkot koostuvat useista eri kerroksista riippuen käytetystä mallista ja muodostetusta arkkitehtuurista. Tavallisin neuroverkkotyyppi on monikerroksinen perseptroniverkko (Multilayer perceptron eli MLP), jossa on vähintään kolme neuronikerrosta: syötekerros (input layer), piilotettu kerros (hidden layer) sekä tuloskerros (output layer)[43]. Syötekerros vastaanottaa annetun syötteen ja kuljettaa sen seuraavalle kerrokselle,

[37] De Bruyne ja Vanleenhove, 2021, s.6.
[38] Helsingin yliopisto, Machine learning: Linear regression.
[39] De Bruyne ja Vanleenhove, 2021, s.6.
[40] De Bruyne ja Vanleenhove, 2021, s.7.
[41] Helsingin yliopisto, Machine learning: Linear regression.
[42] De Bruyne ja Vanleenhove, 2021, s.7-8 ja Massimiliano ja muut, 2020, s.5.
[43] Wikipedia, Monikerroksinen perseptroniverkko; De Bruyne ja Vanleenhove, 2021, s.7 ja Massimiliano ja muut, 2020, s.6.

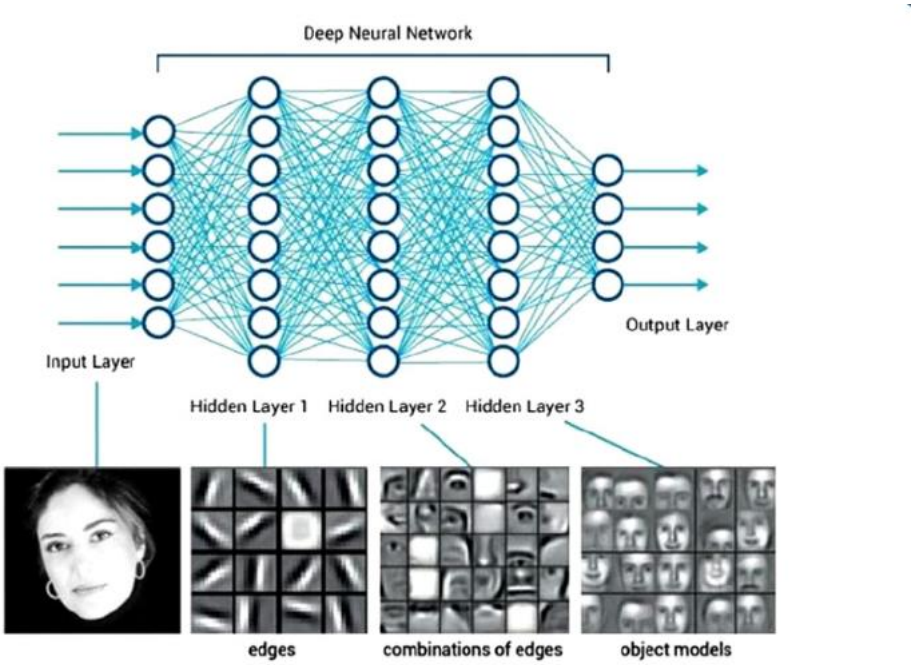
piilotetulle kerrokselle, jossa algoritmi soveltaa parametrejaan (weight ja bias) syötteeseen. Mukautettu syöte kulkeutuu piilotetulta kerrokselta toiselle, kunnes se saavuttaa tuloskerroksen.[44]

Piilotetussa kerroksessa syöte kerrotaan neuronin painoarvolla(weight) ja siihen lisätään vinouma(bias). Vinouma kuvastaa sitä, kuinka kaukana algoritmin ennuste on tarkoitetusta arvosta: matala arvo tarkoittaa, että hermoverkko pystyy antamaan oikeamman tuotoksen. Painoarvo sen sijaan kuvastaa yhteyden vahvuutta eli kuinka paljon vaikutusta muutetulla syötteellä on tuotokseen: korkea arvo tuottaa suurempia muutoksia lopputuloksessa.[45] Piilotetun kerroksen tuloksia eikä vaiheita pystytä päättämään itse[46]. Sekä painoarvo että vinouma ovat molemmat algoritmin opeteltavissa olevia arvoja ja molempien arvot ovat satunnaistettuja ennen opetuksen alkamista[47]. Neuroverkon opettamisvaiheessa tuotosta verrataan ihmisen tarjoamaan kuvaukseen halutusta lopputuloksesta (target)[48] ja sitä mukaan, kun opetusta tapahtuu, parametrien arvot mukautuvat kohti haluttuja arvoja ja oikeaa tuotosta[49]. Kun opetusvaiheen jälkeen neuroverkolle syötetään täysin uusi arvo, tulisi algoritmin pystyä ”ennustamaan” lopputulos oikein[50].

Mikäli neuroverkko sisältää useampia, jopa satoja, piilotettuja ja erikoistuneita kerroksia, kyseessä on syvä neuroverkko eli deep neural network (DNN), jota käytetään syväoppimisessa (deep learning). Geoffrey Hinton keksi pinoa neuroverkkoja päällekkäin ja opettaa jokaiselle tasolle oman osaamisalueen, jolloin jokainen kerros keskittyi tiettyyn toimintoon, esimerkiksi yksi tunnistaisi muotoja, yksi rakenteita ja niin edelleen.[51] Alla esitetyssä kuvassa (kuva 1: kasvojentunnistusjärjestelmä DNN-tekniikalla) on

[44] DeepAI: Weight (Artificial Neural Network).
[45] DeepAI: Weight (Artificial Neural Network).
[46]De Bruyne ja Vanleenhove, 2021, s.7.
[47] DeepAI: Weight (Artificial Neural Network).
[48] Massimiliano ja muut, 2020, s.6.
[49] DeepAI: Weight (Artificial Neural Network).
[50] Massimiliano ja muut, 2020, s.6.
[51] Wikipedia: Deep Learning; De Bruyne ja Vanleenhove, 2021, s.9-10; Siukonen ja Neittaanmäki, 2019, s.76-78 ja Massimiliano ja muut, 2020, s.18.

kasvojentunnistukseen käytettävä syvän neuroverkon arkkitehtuuri. Esimerkissä jokainen kerros oppii eri osiota, muun muassa silmiä tai linjoja, kasvoista.[52]



Kuva 2. kasvojentunnistusjärjestelmä DNN -tekniikalla.

Neuroverkot ovat olleet tähän mennessä paras tapa esimerkiksi luoda hahmon-, kuvan- tai puheentunnistus jopa sille tasolle, että kone suoriutuu ihmistä paremmin kuvien analysoinnista. Neuroverkot vaativat kuitenkin paljon laskentakapasiteettia ja dataa, mistä johtuen vasta viime vuosina neuroverkkoja on pystytty aidosti hyödyntämään.[53] Muun muassa USA:n armeijassa vuonna 2017 on hyödynnetty TAR eli tactical augmented reality- teknologiaa, jonka avulla sotilas saattoi paikallistaa taistelukentällä sijaintinsa ja saada suoraan näkökenttäänsä tietoa lisätyn todellisuuden avulla jopa esteiden taakse nostamalla asettaan, jonka lämpökameraan TAR-laite on yhdistettynä. Vuonna 2018 saatiin neuroverkot aistimaan ihmisen asentoja ja liikkeitä langattomien laitteiden mittaamien radiosignaalien sähkömagneettisen säteilyn avulla. Tätä keksintöä on kutsuttu RF-Pose -teknologiaksi. RF-Pose muuttaa tunnistamansa signaalit tietokoneen ruudulle

[52] Massimiliano, 2020, s.18.
[53] De Bruyne ja Vanleenhove, 2019, s.8-9 ja Massimiliano, 2020, s.18.

liikkuviksi hahmoiksi ja mahdollistaa ihmisen henkilöllisyyden tunnistamisen seinän läpi jopa 83% tarkkuudella.[54]

2.2 Tekoäly lain silmin

Euroopan komission tiedonannossa ”Tekoäly Euroopassa” vuonna 2018 määriteltiin unionin toimesta ensimmäisen kerran tekoäly. Määritelmä oli hyvin lavea ja kuvaili tekoälyä sanoin ”… pyrkii älykkäästi saavuttamaan asetettuja tavoitteita analysoimalla ympäristöään ja toimimalla osittain itsenäisesti. ”[55]. Myöhemmin vuonna 2020 ”White paper on Artificial Intelligence”- julkaisussa, eli Valkoisessa kirjassa tekoälystä, nostettiin painoarvoon teknologinen kehitys, jotta määritelmä joustaisi nopeassa juoksussa ja siitä huolimatta tarjoaisi edelleen oikeudellisen suojan kohteilleen. Lisäksi julkaisu argumentoi tekoälylle tietyt ominaiselementit, joista tekoäly koostuu eli ilman näitä kyseessä ei voisi olla tekoäly: data ja algoritmit. Tässä tapauksessa kuvattiin algoritmeja, jotka on työstetty koneoppimistekniikkaa käyttäen.[56] Kuvaus ei poislukenut muita tekoälytekniikoita, vaan ennemminkin käytti esimerkkinä tällä hetkellä olennaisinta alkeellista tekoälyn muotoa. Tekoälyalgoritmin olennainen ominaisuus tässäkin julkaisussa on oppimiskyky[57]. Tekoäly kykenee oppimaan ja kuitenkin noudattaa sille ennalta määritettyjä tehtäviä itsenäisesti ja ympäristöään tulkiten. Ideana kuitenkin on, että ihminen ei kehittämällä tekoälyä määritä tekoälyn lopullista toimintaa vaan järjestelmä tässä mielessä optimoisi sen, mitä ihminen tahtoo tällä järjestelmällä tehtävän.[58]

Euroopan komission vuonna 2021 antama asetusehdotus tulee, mikäli ehdotus tulee hyväksytyksi, antamaan tekoälylle ”lopullisen” oikeudellisen määritelmän perustan. Tekoälyjärjestelmillä tarkoitetaan säädösehdotuksen mukaan:

---

[54] Siukonen ja Neittaanmäki, 2019, s.70-71.
[55] Euroopan komissio, 2018, s.1.
[56] Euroopan komissio, 2020, s.16.
[57] Euroopan komissio, 2020, s.16.
[58] Euroopan komissio, 2020, s.16.

”a) Koneoppimisen lähestymistavat, mukaan lukien ohjattu, ohjaamaton ja vahvistava oppiminen, jossa käytetään monia erilaisia menetelmiä, kuten syväoppiminen;

b) Logiikkaan ja tietämykseen perustuvat lähestymistavat, mukaan lukien tietämyksen esittäminen, induktiivinen (looginen) ohjelmointi, tietämyskannat, päättelykoneet, (symbolinen) päättely ja asiantuntijajärjestelmät;

c) Tilastolliset lähestymistavat, Bayes-estimointi, haku- ja optimointimenetelmät.”[59]

Säädösehdotuksessa keskitytään keskeisiin toiminnallisuuksiin määrittelemään tekoälyä, esimerkiksi siihen, että tekoäly pystyy tuottamaan ihmisen määrittelemien tavoitteiden perusteella tuloksia kuten ennusteita, sisältöä ja päätöksiä ja samaan aikaan näillä tuloksilla on vaikutuksia ympäristöönsä. Asetuksella pyrkimys on jättää tekoälyn spesifiointi mahdollisimman teknologianeutraaliksi, jotta se ottaisi huomioon teknologian ja markkinoiden nopean kehityksen, mutta samalle se näyttää jättäneen huomiotta oppimiskyvyn osana tekoälyn ominaismääritystä. Tekoälysäädösehdotus ei aseta rajoitusta tekoälyn itsenäisyydelle, vaan muistuttaa, että tekoäly voi olla sulautettu osa tuotetta. Asetuksen myötä komissiolle varataan oikeus laajentaa tietyillä ennalta määritellyillä alueilla käytettävien suuririskisten tekoälyjärjestelmien luetteloa.[60]

Kuvio 1. Euroopan unionin julkaisuja eettiselle tekoälylle.

2.2.1 Tekoälyn etiikka

Järjestelmiin kohdistuu tänä päivänä ja tulevaisuudessa yhä enemmän erityyppistä sääntelyä. Niin järjestelmäkehitykseen, käyttöönottoon ja käyttöön yleisesti sovelletaan ja

[59] Euroopan komissio, 2021a, Liite 1
[60] Euroopan komissio, 2021a, s.13-17 ja johdanto-osan kappale 6.

vaikuttaa useita oikeudellisesti sitovia sääntö. Näitä sääntöjä luodaan Euroopan unionin tasolla, kuin myös kansainvälisellä ja kansallisella tasolla. Merkittävimpiä oikeuslähteitä tässä suhteessa ja soveltamisalassa ovat unionin primaarilainsäädäntö kuten yleinen tietosuoja-asetus eli GDPR, syrjinnän vastaiset direktiivit, konedirektiivi, tuotevastuudirektiivi, asetus maiden muiden kuin henkilötietojen vapaasta liikkuvuudesta, kuluttajalainsäädäntö. Kansainvälisellä tasolla erityisesti YK:n ihmisoikeussopimus ja Euroopan neuvoston yleissopimukset kuten ihmisoikeussopimus saavat luonnollisesti suuren merkityksen.[61]

Tekoälyä käsittelevä korkean tason asiantuntijaryhmän (AI HLEG) ohjeissa ”Luotettavaa tekoälyä koskevat eettiset ohjeet” on todettu etiikan, joka perustuu EU-perussopimuksiin, EU:n perusoikeuskirjaan ja kansainväliseen ihmisoikeuslainsäädäntöön kirjattuihin perusoikeuksiin, olevan parasta tekoälyn etiikkaa[62]. Tekoälyn eettisiin arvoihin kuuluvat muun muassa läpinäkyvyys, oikeudenmukaisuus, hyväntahtoisuus, hyödyllisyys ja kestävyys[63]. Tekoälylle laaditun etiikan avulla voidaan mahdollisesti edistää yksilöiden elämää esimerkiksi elämänlaadun, itsemääräämisoikeuden tai vapauden parantumisen myötä. Vastaisuudessaan, huonosti käsitetyn etiikan kautta tekoäly on potentiaalinen vaikeuttamaan yksilön elämää.[64]

Tekoäly on järjestelmä siinä missä muutkin, mikä tarkoittaa, että myös tekoäly joutuu suhteellisen vahvan sääntelyn, erityisesti positiivisen velvoitteen, alaisuuteen. Näitä tekoälyä koskevia säädöksiä ei voida eikä tulisi tulkita näkökulmana ”Mitä on kiellettyä tehdä” tai ”Mitä ei voida tehdä” vaan, miten aktiivisesti pitäisi tehdä, esimerkkinä henkilötietojen suojaamista edellyttävät toimenpiteet[65]. Lakia yksistään ei kuitenkaan pidetä riittävänä keinona luotettavan tekoälyn kehittämiseksi sikäli, kun lainsäädännön ajantasaisuutta ja muuttumiskykyä ei voida taata[66]. Lainsäädännön jäykkyys ilmenee

[61] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.7-8.
[62] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.11-12.
[63] De Bruyne ja Vanleenhove, 2021, s.54.
[64] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.10.
[65] Alho ja muut, 2018, s. 7 ja Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.8.
[66] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.8.

ongelmallisena eritoten tietotekniikkaan sovellettaessa, koska tietotekninen kehitys kulkee eteenpäin liiankin nopeasti suhteessa oikeusnormiston muuttumisnopeuteen. Näin ollen jokaiseen uuteen tietotekniseen kehitysaskeleeseen ei välttämättä ole siihen soveltuvaa tai tulkittua lainsäädäntöä eikä asioita tule tällöin arvioitua niille kuuluvan todellisen merkityksellisyyden valossa oikea-aikaisesti.

Silmällä pitäen lainsäädännön riittämättömyyttä tekoälyn kehitystä ohjaavana tekijänä Tekoälyä käsittelevä korkean tason asiantuntijaryhmä on luonut luotettavaa tekoälyä koskevia edellytyksiä painopisteinään tekoälyn lainmukaisuus, eettisyys ja luotettavuus sekä luotettavan tekoälyn arviointilistan eli ALTAI:n[67]. Mainitut edellytykset on annettu eettisinä ohjeina tarkoituksenaan edistää ja antaa puitteet luotettavan tekoälyn kehittämiselle ottamatta kantaa näihin varsinaisiin osatekijöihin. Kyse on siis puhtaasti eettisyyteen ja luotettavuuteen johdattelu eikä laillisten oikeuksien antamisesta tai oikeudellisten velvoitteiden asettamisesta. Ohjeet on tarkoitettu kaikille tekoälyalan sidosryhmille, jotka kehittävät, suunnittelevat, ottavat käyttöön, toteuttavat tai käyttävät tekoälyä tai, joihin tekoäly vaikuttaa. Nämä eettiset ohjeet eivät siis rajaudu vain julkiselle sektorille vaan luovat vaikutusta myös yksityisellä puolella yrityksille, yksityishenkilöille, työntekijöille, kuluttajille, organisaatioille, tutkijoille ja niin edelleen. [68]

Tekoäly herättää eettisiä kysymyksiä erityisesti kahdesta syystä; tekoäly on autonominen ja muovautumiskykyinen. Autonomisuudella viitataan kyvykkyyteen toimia kompleksisessa ympäristössä ilman jatkuvaa ihmisen valvontaa ja väliintuloa. Muovautumiskykyisyys sen sijaan viittaa kyvykkyyteen oppia itsenäisesti ympäristöstään kokemusten ja vuorovaikutuksen kautta, ja sen perusteella kyvykkyyteen muuttaa käyttäytymistään.[69] Näihin kahteen ominaisuuteen kietoutuu vahvasti tekoälyn käyttötarkoitus. Sikäli jos tekoälyä käytetään korvaamaan ihmisen toimittamaa päätöksentekoa, tulevat eettisyyden ja luotettavuuden arviointi suurempaan asemaan.

---

[67] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.10 ja Euroopan komissio, 2021b, s.5.
[68] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.2 ja 7-8.
[69] De Bruyne ja Vanleenhove, 2021, s.52.

2.2.2 Luotettavan tekoälyn arviointi (ALTAI)

Eettisellä pohdinnalla ideana on käsitellä kysymyksiä, mitä tekoälyn tulisi ja ei tulisi tehdä, esimerkiksi varmistaa yksilön suojan toteutuminen sekä tunnistaa, mitkä tekijät voivat vaarantaa tätä toteutumasta[70]. Pohdinta kuitenkin vaatii ymmärrystä tekoälyn kehittämisestä, käyttöönotosta ja käytöstä, eli kehityksessä tulisi olla mukana eettistä sekä teknistä näkökulmaa osaava tai osaavia tekijöitä. Pääsääntöisesti ennen tekoälyn käyttöönottoa tulisi varmistua tekoälyn oikeudenmukaisuudesta ja lainmukaisuudesta ja, että sen vaikutukset ovat niiden mukaiset yksilön elämään.[71] Asiantuntijaryhmän eettisten ohjeiden avulla on luotu tekoälyn luotettavuutta ja eettisyyttä koskeva arviointilista, jota voidaan ja tulisi käyttää eräänä arviointimenetelmänä tekoälylle.[72] Lista on myöhemmin julkaistu erillisenä kokonaisuutenansa ”ALTAI” -arviointilistana.

Arviointi jaetaan tapahtuvan viidessä eri kategoriassa: Ihmisen toimijuus ja valvonta; tekninen luotettavuus ja turvallisuus; yksityisyyden suoja ja datahallinto; läpinäkyvyys; monimuotoisuus, syrjimättömyys ja oikeudenmukaisuus; yhteiskunnan ja ympäristön hyvinvointi sekä vastuu. Mikäli tekoäly ohjaa, vaikuttaa tai tukee ihmisen päätöksentekoprosessia esimerkiksi, tekee riskianalyysejä tai riskiennusteita tai jos tekoäly käyttäytyy kuten ihminen tai sillä voi olla vaikutuksia ihmisen kiintymiseen, luottamukseen tai riippumattomuuteen, tulisi ALTAI:n esittämiä kysymyksiä ihmiskeskeisyydestä reflektoida ja pohtia. Huomio kiinnittyy muun muassa siihen, voiko järjestelmä aiheuttaa sekaannusta, onko sen tarjoama päätös, sisältö, neuvo tai muu tuotos tulosta algoritmisesta päätöksenteosta eikä siis toisen ihmisen toimittamasta tuotoksesta.[73] Käyttäjälle tulisi joka tapauksessa ilmoittaa kyseessä olevan tekoälyjärjestelmä, joten myös tähän ilmoitusluontoiseen seikkaan tulisi kiinnittää huomiota. AI HLEG pyrkii viestimään, ettei käyttäjien

[70] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.11 ja De Bruyne ja Vanleenhove, 2021, s.50.
[71] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.11.
[72] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.7-8.
[73] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.7-8.

tulisi antaa liikaa luottaa tekoälyjärjestelmään ja tämän välttämiseksi tulisi tehdä toimenpiteitä[74]. Se, simuloiko tekoäly ihmisen kanssa käytyä sosiaalista kanssakäymistä ja voiko kyseessä oleva tekoälyjärjestelmä aiheuttaa kiintymyssuhdetta tai manipuloida ja luoda riippuvaista käyttäytymistä, ovat erityisiä riskejä, joihin lista pyrkii kiinnittämään huomiota. Vastaisuudessaan, jotta näitä tekoälyn luomia riskejä ihmisen autonomialle ja omalle tahdolle valvottaisiin asianmukaisesti, asettaa arviointilista erityisiä vaatimuksia tällekin. Kysymykset valvonnasta sisältävät muun muassa pohdintaa, onko kyseistä tekoälyjärjestelmää valvovalle henkilölle annettu koulutusta, onko ei-toivotulle haitalliselle vaikutukselle torjuntamekanismia sekä, onko tapaa turvallisesti keskeyttää tekoälyjärjestelmän suorittamaa toimintoa.[75]

ALTAI on eritellyt järjestelmää koskevia hallintamekanismeja silmällä pitäen ihmisen suorittaman valvonnan laatua: human in the loop eli HITL, human on the loop eli HOTL, human in command eli HOC. Human in the loop viittaa ihmisvalvojan kyvykkyyteen astua päätöksenteon syklin joka ikisessä vaiheessa väliin, mikäli tarve tulee. Human on the loop sen sijaan viittaa kyvykkyyteen monitoroida toiminnallisuuksia ja astua väliin suunnittelun sykleissä. Human on command kuvastaa kyvykkyyttä valvoa yleistä toiminnallisuutta sisältäen esimerkiksi taloudelliset ja oikeudelliset vaikutukset. HOC-mekanismin mukana tulee kyvykkyys päättää milloin ja miten tekoälyä käytetään eri tilanteissa eli voidaan myös päättää olla käyttämättä tekoälyä, jos ei haluta käyttää algoritmin tuottamaa päätöstä. [76] Ihmisen ensisijaisuus käy ilmi lisäksi komission antamassa säädösehdotuksessa, jossa mainitaan, että suuririskisessä järjestelmässä on oltava sisäänrakennettuja toiminnallisia rajoituksia, joita järjestelmä ei voi ohittaa ja jotka vastaavat ihmiskäyttäjän komentoihin[77].

Toisena elementtinä luotettavan tekoälyn arviointilista tarjoaa teknisen luotettavuuden ja turvallisuuden. Ideaalinen tilanne on palvelu, johon voidaan luottaa eli palvelulla on

[74] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.7-8.
[75] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.7-8.
[76] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.8.
[77] Euroopan komissio, 2021a, johdanto-osan kappale 48.

käyttövarmuutta. Toinen erityinen teknisen luotettavuuden vaatimus on järjestelmän sietokyky, kun se kohtaa muutoksia. Tekoälyn tulisi siis minimoida ja ennalta ehkäistä ei-toivottua ja odottamatonta vahinkoa koko ajan ja myös silloin, kun järjestelmään tehdään toiminnallisia muutoksia kehityksen varrella. Saman tulisi päteä, mikäli toinen tekoälyjärjestelmä tai ihminen vuorovaikuttaa tekoälyjärjestelmän kanssa haitallisella tavalla. Turvallisuuden ja luotettavuuden arviointia varten kysymyspaletti on jaettu osiin: resilienssi tietoturvahyökkäyksiä kohtaan; yleinen turvallisuus; tarkkuus; toimintavarmuus, paluusuunnitelma ja toistettavuus. Resilienssin arviointia varten pohdinta liittyy muun muassa siihen, voisiko tekoälyjärjestelmä aiheuttaa haitallisia, kriittisiä tai vahingollisia seurauksia riskin tai uhan, kuten väärinkäytön tai käyttökatkon, toteutuessa. Tekoälyjärjestelmän tulisi luonnollisesti noudattaa tietoturvastandardeja ja säännöksiä, joten järjestelmän tulee olla niiden osalta varmistettu. Yleistä turvallisuutta arvioidaan muun muassa tarkistamalla, onko riskit ja riskitasot määriteltynä tekoälyjärjestelmän tietyille käyttötapauksille ja onko identifioitu mahdollisia uhkia. Tekoälyjärjestelmän varalle tulisi lisäksi olla mekanismi arvioida, milloin järjestelmä on kokenut sen verran muutoksia, että on aika arvioida teknistä luotettavuutta ja turvallisuutta uudestaan.[78] Myös nämä teknistä luotettavuutta koskevat vaatimukset ovat löydettävissä tekoälysäädösehdotuksesta.[79]

Tekoälyjärjestelmän tarkkuuden aiheuttamia ongelmia arvioidaan esimerkiksi pohtimalla, voisiko liian vähäinen tarkkuus ja virheet johtavat haitallisiin ja vahingollisiin seuraamuksiin. Järjestelmän virheettömyyden kannalta olennaista on tarjota tekoälyalgoritmille ajantasaista, laadukasta ja laajaa dataa opetusvaiheessa. Tämän arvioimiseen ei kuitenkaan ole annettu erityisiä mittareita, vaan arviointia tekevän pitäisikin itse varmistua tällaisista mittareista ja niiden olemassaolosta järjestelmää varten. Viimeisenä teknisen puolen arviointikohteena on toimintavarmuus, paluusuunnitelmat ja toistettavuus. Esimerkkikysymyksinä tälle osa-alueelle ovat: onko tekoälylle määriteltynä prosessi, jolla monitoroidaan, että tekoäly kohtaa tarkoitetun mukaiset tavoitteet? Ja onko tekoälyä

[78] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.9-10.
[79] Euroopan komissio, 2021a, johdanto-osan kappaleet 49-51.

varten verifiointi- ja validointimetodeja ja dokumentaatiota, joilla pystytään varmistumaan luotettavuudesta ja toistettavuudesta?[80]

Eräs yksilön perusoikeuksista on yksityisyyden suoja ja tähän linkittyen henkilötietojen suoja. Tekoälyjärjestelmän kerätessä dataa henkilöistä tai käyttäessään opetusvaiheessa tällaista datajoukkoa, tulee tekoäly vaikuttamaan automaattisesti yksityisyyden suojaan. Datan turvaamiseksi ja tietosuojaloukkausten ehkäisemiseksi, pyydetään tekoälyn arviointilistassa tarkastelemaan, millaisia vaikutuksia järjestelmällä on tietosuojaan ja perusoikeuksiin, samoin kuin käymään läpi käyttöön otetut mekanismit suojan säilymiseksi[81]. Osa tietosuojaa, ja erityisesti ennalta ehkäisevänä tekijänä, on kunnollinen tiedonhallinta, joka samalla takaa datan laadun ja eheyden. Mietittäessä datanhallintaa, tulee huomiota kiinnittää sisältääkö opetusdata itsessään henkilötietoa ja jos sisältää, millaista. Yleinen tietosuoja-asetus asettaa tiettyjä toimenpidevaatimuksia, jotka pätevät luonnollisesti myös tekoälyjärjestelmien kohdalla.[82] Esimerkiksi DPIA:n eli tietosuojaa koskeva vaikutustenarvion toteuttaminen sekä oletusarvoinen tietosuoja tulevat kysymykseen.

Neljäntenä arviointilistan elementtinä läpinäkyvyys pitää sisällään jäljitettävyyden, selitettävyyden ja viestinnän järjestelmän rajoitteista. Jäljitettävyyden kannalta arvioidaan, dokumentoitiinko järjestelmän kehityksen aikana tarpeellinen tieto, kuten dataan ja päätöksentekoon liittyvät prosessit, asiallisesti. Esimerkiksi tulisi pystyä selvittämään, mitä dataa on käytetty opettamaan tekoälyä antamaan tietty lopputulos käyttäjälleen sekä mitkä säännöt ja mallit johtavat tiettyyn tulokseen. Arviointilistassa lisäksi viitataan tarpeeseen jatkuvasti seurata ja arvioida tekoälyn antaman tuotoksen laatua.[83]

Selitettävyydellä viitataan kykyyn selittää tekniset prosessit sekä tekoälyn tuotosten takana piilevä päättely. Tekoälyn antamat päätökset ja ennusteet tulisi olla selitettävissä ymmärrettävällä tavalla niille, joihin ne luovat myös epäsuoria vaikutuksia. Erityisesti

[80] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.10-11.
[81] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.12.
[82] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.12.
[83] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.14.

seikat, miksi järjestelmä päätyi kyseiseen lopputulokseen ja mitkä syötteet ovat vaikuttaneet tähän. Mikäli näille ei ole mahdollista antaa selitystä, voidaan todeta kyseessä olevan ”musta laatikko”. Mustan laatikon tapauksessa tulisi kuitenkin myös pystyä antamaan selitys annetusta päätöksestä, joten tarpeeseen voi tulla etsiä toinen tapa antaa selitystä. Se, millä tavalla selitys annetaan, riippuu tapauksen kontekstista ja siitä, millä voimakkuudella virheellinen tai epätarkka tuotos vaikuttaa ihmiselämään. Tekoälyä arvioivan osapuolen tulee pohtia, onko tekoälyn päätökset selitettynä käyttäjilleen sekä pystytäänkö seuraamaan jatkuvasti, että käyttäjät ymmärtävät tekoälyn antamat päätökset.[84]

Kommunikointi osana läpinäkyvyyttä vaatii, että tekoälyjärjestelmän rajoitteet ja kyvykkyydet, kuten täsmällisyys, viestitään käyttäjilleen sopivalla tavalla. Esimerkiksi jos kyseessä on interaktiivinen järjestelmä, kuten chatbot, kerrotaanko käyttäjille, että he eivät ole tekemisissä ihmisen kanssa vaan tekoälyn. Luotettavuutta arvioitaessa kommunikoinnin kannalta, lisäkysymyksiä herää, kuvaillaanko järjestelmän hyötyjä, mahdollisia riskejä ja ohjeita sen käyttöön.[85]

Viidentenä osana tekoälyn luotettavuutena tarkastellaan monimuotoisuutta, syrjimättömyyttä ja oikeudenmukaisuutta. Tekoälyn tulee koko sen elinkaaren ajan tukea inkluusiota ja monimuotoisuutta sekä pyrkiä välttämään vääristymiä ja vajavaisuutta. Vääristymät saattaisivat johtaa tiettyihin ryhmiin kohdistuviin ennakkoluuloihin ja diskriminointiin. Vahinkoa voi lisäksi syntyä kuluttajia kohtaan sekä kilpailussa markkinoilla. Järjestelmää tulee siis voida käyttää kaikki riippumatta iästä, sukupuolesta, kyvyistä tai muista ominaisuuksista. Erityisen tärkeään asemaan nousee huomioida saavutettavuus vammaisten näkökulmasta.[86] Käytännössä järjestelmää kehittäessä parhaana lähtökohtana pidetään ”Universal design” -periaatetta, joka korostaa käyttökokemusta lisäämään järjestelmän saavutettavuutta ja käytettävyyttä parhaalla mahdollisella tavalla ja

[84] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.14-15.
[85] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.15.
[86] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.16-17.

mahdollisimman laajasti eri käyttötapauksissa[87]. ALTAI:n kysymyslista keskittyy muun muassa siihen, onko järjestelmän käyttöliittymä sellainen, jota erityistarpeita tai vammoja omaavat henkilöt pystyvät käyttämään eli täyttääkö se esimerkiksi saavutettavuuden ruudunlukuohjelmallakin. Vääristymien kohdalla sen sijaan pohditaan, edustavatko tekoälyn datajoukot tarpeeksi loppukäyttäjiä ja onko järjestelmä testattu erityisiä ryhmiä kattavilla käyttötapauksilla. [88]

Tekoäly voi vaikuttaa ihmiselämän jokaisella alueella, kuten koulutus ja työ, vaikuttaen esimerkiksi sosiaalisiin suhteisiin, kiintymystapoihin ja yleisesti sosiaalisiin taitoihin. Vaikutus voi tietenkin olla positiivinen tai negatiivinen. Jälkimmäistä välttäen tekoälyjärjestelmiä neuvotaan monitoroimaan ja yleisesti AI HLEG- ryhmä muistuttaa, että tekoälyn tulisi hyödyttää kaikkia ihmisiä, myös tulevaisuuden sukupolvia, eikä se missään nimessä saa haitata demokratian toteutumista. Listan mukaan vaikutuksia arvioidaan ympäristön hyvinvoinnin, työn ja osaamisen sekä yhteiskunnan ja demokratian kannalta. Ei riitä, että tekoälyä käytetään ratkomaan ongelmia kuten ilmastonmuutos vaan koko tekoälyn kehityskaaren ajan huomiota tulisi kiinnittää kestävyyteen ja ympäristöystävällisyyteen. Niinpä pohdinta ympäristön osin kohdistuu nimenomaisesti, millaisia vaikutuksia tekoälyllä on ympäristöön ja onko mahdollista luoda mekanismi, jolla valvotaan ympäristövaikutuksia kehityksen ja käytön aikana.[89]

Ympäristöystävällisyyden lisäksi tekoälystä rohkaistaan tekemään tekniikka, joka tukee ihmistä ja luo merkityksellisyyttä työympäristössä. ALTAI:n kysymysluettelo viittaa, että tekoäly tulisi esitellä tarkoituksenmukaisesti työorganisaatiossa sekä sen vaikutukset ihmisen tekemään työhön mitata ja ymmärtää. Esimerkiksi tekoälyjärjestelmä saattaa luoda riskin työvoiman osaamisen ja kyvykkyyden heikentymisestä, jolloin tätä nimenomaista riskiä välttäen tulisi ryhtyä vastatoimenpiteisiin. Yhtä lailla yhteiskunnallisten vaikutusten ja demokratian kohdalla syntyviä riskejä tulisi tunnistaa ja minimoida.

[87] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.29.
[88] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.16-17.
[89] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.19-20.

Esimerkiksi tekoäly saattaa vaikuttaa poliittiseen päätöksentekoon muun muassa kohdistamalla käyttäjälleen väärennettyjä uutisia, erottelemalla ja korostamalla vaaliehdokkaita ja edistämällä totalitaristista käyttäytymistä.[90]

Viimeisenä elementtinä luotettavaa tekoälyä koskevalla arviointilistalla vastuu korostaa tärkeyttä varmistaa kunkin kehityksen ja käyttöönoton osapuolen vastuullisuutta. Tämä osa-alue liittyy riskien hallintaan, mitigointiin sekä auditointiin. Auditoitavuuden kohdalla arvioidaan mahdollisuutta suorittaa arviointeja ja päästä käsiksi dokumentaatioon. Sikäli kun kyseessä on järjestelmä, joka vaikuttaa perus- ja ihmisoikeuksiin, pitää järjestelmä pystyä auditoimaan riippumattomasti kolmansien osapuolten toimesta. Tekoälyä kehitettäessä, samoin kuin käytön aikana, tulisi olla tunnistettuna, arvioituna, dokumentoituna sekä minimoituna potentiaaliset negatiiviset vaikutukset. Kompromisseja saatetaan kuitenkin joutua tekemään näiden vaatimusten yhteydessä, minkä takia tekoälyjärjestelmän olennaisimmat intressit ja arvot pitäisi tunnistaa. Kompromisseja tehdessä tulee aina tunnistaa ja arvioida riskit luonteensa perusteella suhteessa turvallisuuteen ja tekoälyn eettisiin periaatteisiin sisältäen perus- ja ihmisoikeudet. Myös tehdyt kompromissit tulee dokumentoida hyvin.[91]

Ohjeiden tarjoama lista kannanotoista ja kysymyksistä ei ole tyhjentävä vaan luotettavan tekoälyn arvioinnissa tulee huomioida myös ne kysymykset, joita ei kyseisissä ohjeissa ole esitettynä. Näin ollen asiantuntijaryhmä on myös antanut neuvon laatia tekoälyjärjestelmän suunnittelun, käyttöönoton ja käytön eri vaiheissa luettelon täydentävistä kysymyksistä arvioimaan tekoälyn luotettavuutta sekä mukauttamaan tätä kysymysluetteloa, kun kohdataan uusia muuttujia tai siirrytään uuteen kehitysvaiheeseen.[92] Ottaen huomioon koko oikeusjärjestyksen ja yhteiskunnan perustuvan yhteisesti hyväksyttyyn normistomaailmaan, voidaan argumentoida eettisten ohjeiden ja normien olevan yksi

[90] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.20.
[91] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2020, s.21.
[92] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.3 ja 5.

tekoälykehityksen olennaisimmista osa-alueista yhdessä lainsäädännöllisten velvoitteiden noudattamisen kanssa.

2.2.3 Ihmiskeskeisyys tekoälyn periaatteena

Tekoälylle laaditut edellytykset voidaan aiemmissa kappaleissa todetusti jakaa kolmeen osaan: lainmukaisuus, eettisyys sekä tekninen luotettavuus. Lainmukaisuus tässä viittaa kaikkien käyttötapaukseen sovellettavaksi tulevien lakien ja määräysten noudattamista. Eettisyyden kanssa limittyvät lainmukaisuuteen laskettavat tietyt ohjenuorat ovat saaneet erityismaininnan tärkeydessään tekoälyä rakennettaessa: yksilön itsemääräämisoikeuden kunnioittaminen, vahinkojen välttäminen, oikeudenmukaisuus ja selitettävyys.[93] Erityisesti selitettävyys on tekoälyssä haluttu ominaisuus, koska jokaisella on oikeus saada perustelut häntä itseään koskevalle oikeudellisesti vaikuttavalle päätökselle. Tekoälyä käytettäessä näin ei aina voida todeta tapahtuvan. Esimerkiksi ymmärrys neuroverkkojen avulla oppivasta tekoälystä ei riitä, jotta pystyttäisiin aina selittämään, miten tuo tekoäly päätyi tarjoamaansa ulosantiin[94].

Avainsanana eettisyydelle tekoälyyn yhdistettynä on *ihmiskeskeisyys*, jonka itsessään on tarkoitus heijastaa sovittujen perus- ja ihmisoikeuksien toteutumista[95]. Käytännön tasolla tämä tarkoittaa, että

> *”ihmisellä on ainutlaatuinen ja jakamaton moraalinen ensisijaisuus kansalaiselämässä, politiikassa, taloudessa ja yhteiskunnassa”.*[96]

Tekoälyn olisi oltava ensi sijassa työväline ihmisille ja hyvää tuottava voima yhteiskunnassa perimmäisenä tavoitteenaan ihmisten hyvinvoinnin lisääminen[97].

---

[93] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.2.
[94] Topol, 2019, s.94.
[95] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.12 ja Euroopan komissio, 2021b, s.1.
[96] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.12.
[97] Euroopan komissio, 2021a, s.1 ja Euroopan komissio, 2021b, s.9.

Ihmiskeskeisyyden varmistamiseksi Euroopan komission julkaisussa ”Coordinated Plan on Artificial Intelligence 2021” esiteltiin neljä avainehdotusta, joilla tekoälystä saadaan ihmiskeskeistä sekä luotettavaa, turvallista ja kestävää. Nämä ehdotukset korostivat muun muassa keskitettyä säädäntöä, jaettua dataa sekä yhteistyötä, viitekehysten luontia ja rahoitusmahdollisuuksia. Myös tässä yhteydessä esille noussut seikka oli tekoälyn toiminta ihmisiä varten. [98] Ihmiskeskeisyyden varmistamiseksi komissio esitti pyrkimykseksi edistää yksilöinen osaamista ja ymmärrystä tekoälystä ja sen tekniikoista ulottuen niin ikään teknisen näkökulman ulkopuolelle[99].

Tietenkin perus- ja ihmisoikeudet tulevat joka tapauksessa oikeudellisesti sitoviksi kriteereiksi kaikissa niiden soveltamisalaan kuuluvissa tilanteissa myös tekoälyn kohdalla, mutta esimerkiksi Euroopan unionin perusoikeuskirjan soveltamisala rajoittuu vain EU-oikeuden aloihin. Avuksi tässäkin tilanteessa tulee kansainvälinen ihmisoikeuslainsäädäntö ja Euroopan ihmisoikeussopimus, jotka sitovat Euroopan unionin jäsenvaltiota myös aloilla, jotka eivät kuulu EU-oikeuden soveltamisalaan. Täten ei kuitenkaan voida todeta perus- ja ihmisoikeuksien ulottuvan täydellisesti joka paikkaan kaikessa laajuudessaan. [100]

Ihmisarvon kunnioitukseen sisältyy ajatus, ettei muun ihmisen lisäksi tekoälyjärjestelmän kaltainen teknologia saa pienentää, vaarantaa tai tukahduttaa kenenkään yksilön itseisarvoa. Tekoälyn olisi siis asiantuntijaryhmän eettisten ohjeiden mukaan kohdeltava kaikkia ihmisiä kunnioittavasti moraalisina toimijoina eikä sellaisina kohteina, joita seulotaan, lajitellaan, pisteytetään, holhotaan, ehdollistetaan tai manipuloidaan. Sen sijaan tekoälyä kehitettäessä tarkoitus on edistää ja suojella ihmisen fyysistä ja henkistä koskemattomuutta. Yksilön olisi myös voitava tehdä elämäänsä koskevat päätökset, mikä edellyttää valtiollisten tai valtiosta riippumattomien organisaatioiden toimia yksilön

[98] Euroopan komissio, 2021b, s.3-4.
[99] Euroopan komissio, 2021b, s.3-4.
[100] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.12.

vapauden turvaamiseksi. Näin ollen jokaisen tekoälyä kehittävän tulisi pyrkiä jopa estämään pakottaminen, tarpeeton valvonta, harhaanjohtaminen sekä vilpillinen manipulointi.[101]

Huomiota pyydetään kiinnittävän niissä tilanteissa, joihin liittyy haavoittuvia ryhmiä kuten lapsia, vammaisia tai periaatteellisesti heikommassa asemassa tai syrjäytymisvaarassa olevia sekä tilanteissa, joissa vallitsee vallan ja tiedon epätasapaino kuten työnantaja-työntekijä-suhde tai kuluttajan ja yrityksen välinen suhde[102]. Tekoälyjärjestelmän olisi varmistettava kaikkien yksilöiden yhtäläinen ihmisarvon kunnioitus kuitenkin niin, että sallitaan erilaisten tilanteiden erottelu objektiivisesti perusteltuina. Järjestelmä ei saisi tuottaa epäoikeudenmukaisesti puolueellisia tuotoksia.[103] Lisäksi tulisi tunnustaa tekoälyyn liittyvät riskit ja negatiiviset vaikutukset, joita itsessään voi kuitenkin olla hyvin vaikea ennakoida, tunnistaa ja mitata, esimerkkinä demokratiaan, oikeusvaltioperiaatteeseen ja oikeudenmukaiseen jakautumiseen tai ihmismieleen liittyvät vaikutukset[104]. Tekoälyjärjestelmän on palveltava demokratian säilymistä ja edistämistä eikä vastaisuudessaan saa heikentää demokraattisia prosesseja, ihmisten suorittamaa harkintaa tai demokraattista vaalijärjestelmää[105].

2.2.4 Suuririskiset tekoälyjärjestelmät

Vuonna 2020 Euroopan komissio julkaisi tekoälyä koskevat toimintapoliittiset suuntaviivat julkaisussa ” White Paper on Artificial Intelligence: an European approach to excellence and trust”. Tämän politiikan aikomus on ollut edistää tekoälyn käyttöönottoa ja puuttua tekoälyteknologiaan liittyviin riskeihin määritellyissä rajoissa. Jatkumona edellä mainitulle julkaisulle komissio antoi vuonna 2021 ehdotuksen ensimmäiseksi tekoälysäädökseksi. Kyseessä on ”Ehdotus Euroopan parlamentin ja neuvoston asetukseksi tekoälyä koskevista yhdenmukaistetuista säännöistä (tekoälysäädös) ja tiettyjen unionin

[101] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.12-13.
[102] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.2.
[103] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.13.
[104] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.2.
[105] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.13.

säädösten muuttamisesta 2021/0106 (COD)". Säädösehdotuksen nimenomainen tarkoitus on toteuttaa valkoisen kirjan jälkimmäinen tavoite puuttua tekoälyn riskeihin tiettyjen säädöksessä nimettyjen rajojen sisällä. Säädösehdotuksen itsenäiset tavoitteet ovat seuraavat: tekoälyjärjestelmät noudattavat perusoikeuksia ja unionin arvoja; taataan oikeusvarmuus; tehostetaan tekoälyjärjestelmiin sovellettavia perusoikeuksia ja turvallisuusvaatimuksia koskevan lainsäädännön hallinnointia ja täytäntöönpanoa; helpotetaan sisämarkkinoiden kehittämistä tekoälysovelluksille.[106]

Tekoälyjärjestelmästä perusoikeuksille aiheutuva haitallisuus on erityisen merkityksellinen silloin, kun järjestelmä luokitellaan suuririskiseksi. Esimerkiksi, jos tekoälyjärjestelmää käytettäisiin määrittämään etuuksien ja palveluiden epäämisestä, rajoittamisesta, peruuttamisesta tai takaisinperimisestä, järjestelmän tekemä päätös vaikuttaisi henkilön toimeentuloon ja täten hänen sosiaaliseen suojeluunsa ja mahdollisesti syrjimättömyyteen sekä ihmisarvoon.[107] Myös oikeusviranomaisia tosiseikkoja ja lainsäädäntöä tutkimaan ja tulkitsemaan auttavat sekä lainsäädäntöä konkreettisiin tosiseikkoihin soveltavat järjestelmät olisi tarpeen luokitella suuririskisiksi vinoumien, virheiden ja läpinäkymättömyyden ehkäisemiseksi. Hallinnollisiin tukitoimiin, esimerkiksi tietojen pseudonymisointiin, käytettävät oikeusviranomaisten järjestelmät eivät olisi suuririskisiä tekoälyjärjestelmiä.[108] Sen sijaan esimerkiksi vero- ja tulliviranomaisten tekoälyjärjestelmiä ei olisi tarkoitus luokitella suuririskisiksi tekoälyjärjestelmiksi, kun järjestelmää käytetään rikosten ennalta ehkäisemiseen, tutkimiseen, paljastamiseen tai rikoksiin liittyviin syytetoimiin[109]. Lisäksi kokonaan säädöksen soveltamisalan ulkopuolelle jäisi järjestelmät, jotka on tarkoitettu käytettäväksi yksinomaisesti sotilaallisiin tarkoituksiin[110].

Tekoälysäädös perustuu riskiperusteiseen sääntelymalliin tarkoittaen, että oikeudelliset toimenpiteet on määritelty nimenomaisesti tähän tapaukseen sopiviksi perustuen esille

[106] Euroopan komissio, 2021a, s.1-3.
[107] Euroopan komissio, 2021a, johdanto-osan kappale 28 ja 37.
[108] Euroopan komissio, 2021a, johdanto-osan kappale 40.
[109] Euroopan komissio, 2021a, johdanto-osan kappale 38.
[110] Euroopan komissio, 2021a, johdanto-osan kappale 12.

nouseviin perusteltuihin huolta aiheuttaviin tilanteisiin. Tässä tapauksessa erotellaan toisistaan tekoälyn käyttö, joka aiheuttaa riskin, joka ei ole hyväksyttävä; suuren riskin; tai vähäisen tai minimaalisen riskin. Rasitetta aiheutetaan lähinnä suuririskisille tekoälyjärjestelmille, ja muille vähäriskisemmille syntyy erityisiä läpinäkyvyysvelvoitteita muun muassa ilmoittaa tekoälyn käytöstä, kun järjestelmä on vuorovaikutuksessa ihmisen kanssa. Suuririskisille järjestelmille asetetut vaatimukset koskevat dataa, dokumentointia ja jäljitettävyyttä, läpinäkyvyyttä, ihmisen suorittamaa valvontaa sekä tarkkuutta ja varmuutta.[111]

Asetuksessa on annettu lista kielletyistä käytännöistä, jotka koskevat kaikkia niitä tekoälyjärjestelmiä, jotka ovat Euroopan unionin arvojen vastaisia eli esimerkiksi loukkaa perusoikeuksia. Kiellon piiriin kuuluu muun muassa *merkittävä* mahdollisuus *manipuloida* ihmisiä tekniikoilla, joista he eivät ole *tietoisia*; *haavoittuvien* ryhmien, kuten lasten, hyväksikäyttö niin, että heidän käyttäytymistään vääristetään lisäten todennäköisyyttä aiheuttaa *heille tai toiselle psyykkistä tai ruumiillista haittaa.* Tällaisten tekoälyjärjestelmien markkinoille saattaminen, käyttöönotto tai käyttö olisi siis kiellettävä. Toisekseen kiellettävän halutaan tekoälyjärjestelmät, joita voidaan käyttää *pisteyttämään luonnollisia* henkilöitä *yleistä* tarkoitusta varten, koska ne voivat syrjivien tulosten seurauksena johtaa syrjimättömyyden ja tasa-arvon loukkaukseen.[112] Kolmantena kiellon kohteena on *luonnollisen* henkilön *reaaliaikainen* biometrinen etätunnistus *julkisissa* tilanteissa *lainvalvontatarkoituksessa,* koska se voi vaikuttaa yksityiselämään negatiivisesti. Tämän perusteella etätunnistusta voidaan mainitussa tilanteessa toteuttaa jälkikäteen kuitenkin jo kerätyn materiaalin, esimerkiksi videokameran nauhoitteen, turvin. Mainittuja reaaliaikaisia etätunnistusjärjestelmiä voitaisiin kuitenkin käyttää kolmessa *tyhjentävässä* tilanteessa: rikoksen uhrien ja kadonneiden lasten etsintä; luonnollisen henkilön henkeen tai fyysiseen turvallisuuteen kohdistuvan uhan tai terrori-iskun uhat; sekä

[111] Euroopan komissio, 2021a, s.3-7, 13 ja johdanto-osan kappale 14.
[112] Euroopan komissio, 2021a, s.3-7, 13 ja johdanto-osan kappaleet 16 ja 17.

neuvoston päätöksessä 2002/584/YOS tarkoitettujen rikosten tekijöiden tai epäiltyjen havaitsemiseen, paikantamiseen, tunnistamiseen tai syytteeseenpanoon.[113]

Suuririskiset eli *merkittävästi* yksilön *terveyttä*, *turvallisuutta* tai *perusoikeuksia* vaarantavat järjestelmät ovat sallittuja ja Euroopan markkinoille julkaistavia, mikäli ne täyttävät määritellyt pakolliset vaatimukset ja niille tehdään vaatimustenmukaisuuden ennakkoarviointi. Asetuksen edellä määrittelemien ehtojen alla tekoälyjärjestelmän kategorisointi ei voisi perustua järjestelmän toiminnallisuuksiin vaan myös sen käyttökohteeseen ja -tarkoitukseen.[114] Valkoinen kirja tekoälystä nosti esille kaksi huomioitavaa kriteeriä, jotka sittemmin tulivat osaksi asetusehdotusta. Kriteereistä ensimmäisen oli tarkoitus rajata lainkäytön väliintulo sellaiselle alueelle, jolla voidaan yleisesti kuvata riskien olevan todennäköisiä, esimerkiksi terveydenhuolto. Toinen kriteeri taasen kiinnitti huomiota riskin laatuun, sillä kaikki riskialueella käytetyt järjestelmät eivät käyttönsä perusteella voisi luoda tarkoitettua riskiä, esimerkiksi terveydenhuollon ajanvarausjärjestelmä. Tässä tilanteessa riskin laatua arvioitiin muun muassa, voiko järjestelmä luoda oikeudellisia vaikutuksia, aiheuttaa fyysistä vammaa, kuolemaa tai merkittävää materiaalista tai immateriaalista vahinkoa. Edellä kuvattujen kahden kriteerin avulla pystyttäisiin kohdentamaan ja rajaamaan tulevaisuudessa säädösten alaisuuteen joutuvia tekoälyratkaisuja.[115]

Vuonna 2021 annetussa asetuksessa määriteltiin valkoista kirjaa tarkemmin, mitkä alat joutuvat sääntelyn kohteeksi: mikäli käyttö limittyy koskien terveyttä, turvallisuutta tai perusoikeuksia, joutuu tekoälyjärjestelmä suuririskisyyden arviointikohteeksi. Toisena arviointiin vaikuttavana tekijänä olisi edelleen mahdollisen riskin vakavuus ja esiintymistodennäköisyys. Edellä todetuin perustein, biometriset etätunnistusjärjestelmät olisivat suuririskisiä järjestelmiä eivätkä sellaisinaan ole kiellettyjä. Muita suuririskisiä järjestelmiä olisivat muun muassa tekoälylliset turvakomponentit tieliikenteen sekä vesi-, kaasu-,

[113] Euroopan komissio, 2021a, johdanto-osan kappaleet 18 ja 19.
[114] Euroopan komissio, 2021a, s.14 ja johdanto-osan kappale 27.
[115] Euroopan komissio, 2020, s.17.

lämmitys- ja sähkönhuollon hallinnassa ja toiminnassa sekä esimerkiksi järjestelmä, joka määrittää yksilön pääsyn oppilaitokseen tai päättää yksilön rekrytoimisesta tai työsuhteen lopettamisesta tai jota käytetään määriteltäessä yksilön oikeutta saada etuuksia tai luottoa.[116] Listaus suuririskisistä tekoälyjärjestelmistä on annettu säädösehdotuksen liitteessä III, jonka muuttamiseksi Euroopan komissiolle on annettu valta artiklan 73 mukaisesti.

Suuririskisiin tekoälyjärjestelmiin liittyvät vaatimukset koskisivat erityisesti datan laatua, teknistä dokumentaatiota, tietojen säilyttämistä, läpinäkyvyyttä ja luovuttamista, ihmisen suorittamaa valvontaa ja kyberturvallisuutta[117]. Erityiset oikeudelliset vaatimukset suuririskisiä tekoälyjärjestelmiä kohtaan on määritelty tekoälysäädösehdotuksen toisessa luvussa: data ja sen hallinta, dokumentointi ja tietojen säilyttäminen, läpinäkyvyys ja tietojen antaminen käyttäjille, ihmisen suorittama valvonta, luotettavuus, tarkkuus ja turvallisuus. Kolmas luku asettaa tekoälyn tarjoajille ja käyttäjille velvoitteita.[118]

[116] Euroopan komissio, 2021a, johdanto-osan kappaleet 33-37.
[117] Euroopan komissio, 2021a, johdanto-osan kappale 43 ja Euroopan komissio, 2020, s.18.
[118] Euroopan komissio 2021a, s.14-15.

3 Tekoälyn vastuuasettelu

3.1 Vastuu tekoälyn tekijällä

Komission antamassa tekoälysäädöksen ehdotuksessa ilmaistaan osana syrjimättömyyttä olevan vaatimukset erityisesti datajoukkoja kohtaan[119]. Käytännössä liian suppea ja pieni joukko opetusdataa johtaa helposti tekoälyn antaman tuotoksen vääristymiin. Esimerkiksi Nikon kamerassa olleessa tunnistuksessa suljetuille silmille, kamera tulkitsi virheellisesti aasialaisten pitävän silmiä kiinni, koska opetusdata sisälsi vain kaukasialaisia ihmisiä. Tässä tapauksessa tekoäly ei oppinut kokonaisuudesta, joka kattaisi eri käyttäjät, joten malli ei pystynyt yleistämään oikein opetusdatan malleja ja vertaamaan uuteen dataan. Sikäli jos dataa ei kerätä samoin eri ihmisryhmistä, esimerkiksi vähemmistöistä saattaa olla vähemmän dataa saatavilla, voivat henkilöt joutua aliedustetuiksi opetusmateriaaleissa. [120]

Toisekseen datan laatu tulee ratkaisevana tekijänä ehkäisemään vääristymiä tekoälyn ulosannissa. Mikäli tekoälyä opetettaisiin jo vääristyneellä tai epäfaktuaalisella datalla, tekoäly alkaisi soveltamaan tätä opittua vääristymää kaikkeen uuteen dataan, jolloin myös tuotos olisi vääristynyt. Esimerkiksi Amazon on käyttänyt algoritmeja, jotka opetettiin etsimään toistuvia kaavoja hakemuksista 10 vuoden ajalta ja sitten soveltamaan opittua uusiin hakemuksiin. Lopputuloksena algoritmi alkoi rankaisemaan hakemuksia, jotka sisälsivät sanan ”nainen”. Toisessa tapauksessa eräässä yliopistossa luotettiin tietokoneen suorittamaan samankaltaiseen rankkaukseen hakemuksissa. Opetusdatana oli aiemmin hyväksytyt hakemukset ja lopputuloksena algoritmi alkoi syrjimään täysin päteviä naisia. On myös täysin mahdollista, että opetusdatan luoneet henkilöt ovat itsessään diskriminoineet, esimerkiksi sen sijaan, että tarkistaisivat luottohistorioita ja ammatillista menestystä, henkilöt olisivatkin arvioineet kohteen ihonväriin perustuen.[121]

---

[119] Euroopan komissio, 2021a, s.4.
[120] De Bruyne & Vanleenhove, 2021, s. 10, 86-87 ja 139-140.
[121] De Bruyne & Vanleenhove, 2021, s. 138-139 ja 144.

Vaatimukset asetetaan tekoälysäädösehdotuksessa koskien nimenomaan suuririskisiä tekoälyjärjestelmiä. Ehdotus painottaa datan laatua erityisesti sellaisessa tilanteessa, jossa käytetään tekniikoita, jotka sisältävät kouluttamista, diskriminoinnin ehkäisemiseksi. Datajoukkojen olisi oltava ehdotuksen mukaan:

> *"- - merkityksellisiä, edustavia ja virheettömiä ja täydellisiä järjestelmän suunniteltuun käyttötarkoitukseen nähden. Niillä olisi lisäksi oltava asianmukaiset tilastolliset ominaisuudet, myös niiden henkilöiden tai henkilöryhmien osalta, joihin suuririskistä tekoälyjärjestelmää on tarkoitus käyttää. - - olisi erityisesti otettava niiden käyttötarkoituksen edellyttämässä laajuudessa huomioon ne ominaispiirteet, ominaisuudet tai osatekijät, jotka ovat ominaisia sille maantieteelliselle, käyttäytymiseen liittyvälle tai toiminnalliselle ympäristölle tai asiayhteydelle, jossa suuririskistä tekoälyjärjestelmää on tarkoitus käyttää.".* [122]

Jotta pystyttäisiin varmistamaan syrjimätön, luotettava ja vastuullinen pääsy datajoukkoihin, EU:n aikomus on muodostaa yhteinen data-alusta sen sijaan, että data olisi sijoittuneena instituutioille, jäsenvaltioille ja yrityksille itselleen[123]. Kyseessä on Euroopan unionin yhteinen data-avaruus, josta kerron vielä myöhemmin enemmän. Yhteistä data-avaruutta ja datan uudelleenkäyttöä sääntelevät erityisesti Euroopan parlamentin ja neuvoston asetus (EU) 2022/868 eurooppalaisen datan hallinnoinnista ja asetuksen 2018/1724 muuttamisesta (datanhallinta-asetus) sekä Euroopan parlamentin ja neuvoston direktiivi (EU) 2019/1024 avoimesta datasta ja julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä. Direktiivi toteaa sisämarkkinoiden toteuttamisen yhdeksi tärkeimmistä tavoitteista olevan edellytykset edistää palvelujen ja tuotteiden kehittämistä unionissa sekä toteaa käytännössä julkisella intressillä tuotetun palvelun yhteydessä kerätyn, tuotetun, jäljennetyn tai jaetun tiedon olevan tärkeä raaka-aine palveluille esimerkiksi tekoälyn sisällönlähteenä[124]. Tämän direktiivin määrittelemä uudelleenkäytettävä

---

[122] Euroopan komissio, 2021a, johdanto-osan kappale 43 ja 44.
[123] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019a, s.28 ja Euroopan komissio, 2021a, johdanto-osan kappale 45.
[124] Euroopan parlamentin ja neuvoston direktiivi (EU) 2019/1024 avoimesta datasta ja julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä, johdanto-osan kappale 13.

tieto tulisi saataville mahdollisesti reaaliaikaisten API-rajapintojen kautta[125]. Direktiivi ei vaikuta suojaan, joka annetaan henkilötietojen käsittelyn yhteydessä vaan henkilötietoja voidaan uudelleenkäyttää vain GDPR:n vaatimusten mukaisesti, kun tiedot on kerätty:

> *”tiettyä, nimenomaista ja laillista tarkoitusta varten, eikä niitä saa käsitellä myöhemmin näiden tarkoitusten kanssa yhteensopimattomalla tavalla”. (GDPR 5 art. 1 kohta b alakohta).*

sekä käsittely täyttää artiklan 6 periaatteet. Mikäli uudelleenkäytön kohteeksi joutuu henkilötietoja, esimerkiksi terveydenhuoltoalalta, tulee tässä toteuttaa GDPR:n vaatimat tietosuojan vaikutustenarvioinnit.[126]

Edellä mainittu datanhallinta-asetus asettaa datan uudelleenkäytölle ehdoiksi anonymisoinnin, mikäli kyseessä on henkilötiedot ja kaupallisesti luottamuksellisten tietojen, liikesalaisuuksien tai teollis- ja tekijänoikeuksilla suojatun tiedon kohdalla minkä tahansa muun suojamenetelmän (Datanhallinta-asetus (EU) 2022/868, art. 5 kohta 3. alakohta a). Anonyymillä tiedolla tarkoitetaan tietoa, joka ei liity tunnistettuun tai tunnistettavissa olevaan henkilöön vaan on anonymisoitu tunnistamattomaksi[127]. Datanhallinta-asetuksen sama artikla kieltää tunnistamasta uudelleen rekisteröityjä ja asettaa uudelleenkäyttäjälle vaatimuksen toteuttaa tarvittavat tekniset ja operatiiviset toimenpiteet tämän estääkseen.

Valkoinen kirja tekoälystä kehottaa ylläpitämään täsmällistä dokumentaatiota tekoälyn koulutukseen ja testaamiseen käytetyistä datajoukoista sisältäen kuvauksen datan perusominaisuuksista ja millä perustein data valittiin. Lisäksi kehotus neuvoo pitämään dokumentaatiota itse kehityksestä ja opetustavoista, -prosesseista ja -tekniikoista, joilla tekoäly testattiin ja lopulta validoitiin. Joissain tapauksissa olisi myös tarpeellista sisällyttää itse datajoukot dokumentaatioon. Dokumentaation tulisi olla saatavilla pyynnöstä, kun

---

[125] Euroopan parlamentin ja neuvoston direktiivi (EU) 2019/1024 avoimesta datasta ja julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä, johdanto-osan kappale 32.
[126] Euroopan parlamentin ja neuvoston direktiivi (EU) 2019/1024 avoimesta datasta ja julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä, johdanto-osan kappaleet 52 ja 53.
[127] Euroopan parlamentin ja neuvoston direktiivi (EU) 2019/1024 avoimesta datasta ja julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä, johdanto-osan kappale 52.

tekoälyä on tarvetta testata tai tarkistaa tähän valtuutetun toimesta.[128] Verraten komission antaman säädösehdotuksen mukaisesti erityisesti olisi dokumentoitava tietoa, miten suuririskinen tekoälyjärjestelmä on kehitetty ja miten se toimii eri elinkaaren aikoina. Tällaista säilytettävää tietoa olisi yleiset ominaisuudet, valmiudet ja rajoitukset, algoritmit, data, koulutus-, testaus- ja validointiprosessit sekä riskienhallintajärjestelmää koskevat asiakirjat. Dokumentaation olisi aina oltava ajan tasalla.[129]

Kuten edellä kuvatuista vaatimuksista käy ilmi, vastuuta tekoälyn kehittämisestä annetaan ainakin datajoukkojen suunnittelijoille sekä datan käsittelijöille ja uudelleenkäyttäjille. Suorin vaikutus näistä tekoälyn toimintaan on datajoukoilla itsellään eli tekoälyn kehityksessä mukana olevilla tahoilla. AI HLEG kehottaa monitoroimaan tekoälyn vaikutuksia ja varmistamaan tekoälyn oikeudenmukainen toiminta vahingollisia vaikutuksia vastaan[130]. Olennainen tapa varmistua tekoälyn oikeasta tarkoitetusta toiminnasta on järjestelmän testaaminen eri tilanteita ja käyttötapauksia vasten. Tämä nimenomainen vaatimus on kirjattu AI HLEG -ryhmän ohjeisiin sekä komission antamaan tekoälysäädösehdotuksen artiklan 9 alakohdassa 5. Yleisesti pyrkimys olisikin varmistaa tekoälyyn liittyvien eettisten ohjeiden käyttö julkisella sektorilla käytettävien tekoälyjärjestelmien kohdalla[131]. Kyseisillä ohjeilla tarkoitetaan nimenomaan AI HLEG- asiantuntijaryhmän luomia eettisiä ohjeita ”Assesment List for Trustworthy Artificial Intelligence (ALTAI) for self-assesment”, jotka soveltuvat tekoälyn suunnittelijoille tai kehittäjille, datatieteilijöille, spesialisteille, käyttäjille ja niin edelleen, sovellettaviksi[132].

Käytännössä tarve olisi jatkuvasti arvioida, voiko tekoäly luoda riskin, jota ei ole lainsäädännössä huomioituna. Henkilön ei esimerkiksi tulisi joutua kokemaan tarpeetonta henkilökohtaista, fyysistä tai mentaalista jäljitystä tai tunnistusta, profilointia tai tuuppausta sellaisen tekoälyn kautta, joka käyttää biometrisiä tunnistusmekanismeja esimerkiksi

[128] Euroopan komissio, 2020, s.19-20.
[129] Euroopan komissio, 2021a, johdanto-osan kappale 46.
[130] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019a, s.10.
[131] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019a, s.20.
[132] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.4.

tunteiden seurantaa, DNA:ta, silmänliikkeitä tai äänen- tai kasvojentunnistusta.[133] Erityisiä eettisiksi arvoiksi nostettuja vaatimuksia tekoälylle ovat muun muassa ihmisen toimijuus ja ihmisen suorittama valvonta, tekninen luotettavuus ja turvallisuus, yksityisyyden suoja, läpinäkyvyys ja syrjimättömyys[134]. Eettisten ohjeiden lisäksi tekoälyn kehityksessä mukana olevia tahoja sitoo tietenkin myös voimassa oleva kansallinen, kansainvälinen ja EU-lainsäädäntö kuten perustuslaki perusoikeuksineen ja yleinen tietosuoja-asetus henkilötietojen suojaamiseksi.

Tekoälyä kehittävät tahot sekä algoritmi ja data itsessään saattavat olla vahvasti tekoälyn virheeseen osallisia, koska tietokoneohjelmana tekoäly on tulos inhimillisestä luomis- ja kehitystyöstä. Erityisen ongelmallista tilanteessa kuitenkin on, että kehittäneet tahot eivät välttämättä tiedä, miten tekoälyjärjestelmä toteuttaa toimintonsa esimerkiksi neuroverkkojen tapauksessa. Tällaisen tilanteen vallitessa, kun esimerkiksi kehittäjä ei tunne kaikkia järjestelmän toimintaan liittyviä seikkoja, on yleisesti tulkittu tilanteena, joka ei loisi vastuuta esimerkiksi vahingon korvaamiseksi.[135]

3.2 Vahingonkorvausvastuu

Tekoälyjärjestelmän tekemien virheiden vastuuttamisesta on kirjoittanut Ismo Kallioniemi kirjassaan ”Tekoälyoikeus” keskittyen sopimus- ja varallisuusoikeuden näkökulmiin. Kallioniemi hyödyntää tavanomaista tuottamusvastuun ideologiaa sovellettuna tekoälyn tuottamiin vahinkoihin. Korvausvastuun syntymiseen suhteessa sivullistahoon vaatii kolme edellytystä: vahingon aiheutuminen, tuottamuksen olemassaolo tai ankara vastuu, ja toiminnan ja vahingon välinen syy-yhteys[136]. Vahingolla tarkoitetaan haitallista seuraamusta, joka voi olla fyysinen tai taloudellinen. Tuottamuksen tai ankaran vastuun määrittelee harjoitetun toiminnan ominaisluonne. Ankara vastuu johtaa, että toiminnan

[133] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019a, s.40.
[134] Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019b, s.2-3.
[135] Kallioniemi, 2022, s.159-162.
[136] Kallioniemi, 2022, s.140.

harjoittaja on tuottamuksesta riippumatta velvollinen korvaamaan vahinko, jos syy-yhteys tunnistetaan. Tuottamus sen sijaan on huolimattomuutta, ja jos toiminnan luonteeseen asetettua huolellisuustasoa ei saavutettu, syntyy korvausvastuu. Huolellinen toiminta ei siis synnytä tuottamusvastuuta ja täten vahingonkorvausvastuuta. Käytännössä merkitystä ei näissä tapauksissa ole, onko tekoälyä ylipäätänsä käytetty, koska tekoäly ei itsessään ole vahingonkorvausvastuun luova tai vapauttava peruste. Tekoäly voi kuitenkin vaikuttaa huolellisuuden tasoon sekä riskin todennäköisyyteen esimerkiksi vuorovaikutteisuudellaan, samoin tekoälyn autonomian taso voi tulla arvioitavaksi sikäli, kun avustavana järjestelmänä tekoäly toimisi vain työkaluna ja vastaisuudessa täysin autonomisena arvioinnin kohteeksi joutuisi lisäksi tekoälyn ominaisuudet ja sen valvonta. Täysin autonomisen tekoälyllä varustetun järjestelmän tuottamus perustuisi ominaisuuksien riittämättömyyteen nimenomaiseen harjoitettuun toimintaan. Joka tapauksessa toiminnan harjoittaja on vastuussa tehtävän huolellisesta toteuttamisesta ja tämän valvonnasta.[137]

Tekoälyn kolmannelle aiheuttama vahinko voi edellä esitetyn mukaisesti tulla korvattavaksi perustuen tekoälyn käyttäjän toimintaan. Mikäli tekoäly on myyty palveluna tai tuotteena, voitaisiin tässä tapauksessa soveltaa kauppalakia tai kuluttajansuojalakia sekä itse kauppasopimusta. Myyjän korvausvelvollisuus näissä tilanteissa perustuisi tavaran tai palvelun virheeseen ja myyjän olisi suoritettava se, mitä on sovittu sekä korvattava virheestä aiheutunut vahinko (Kauppalaki 1987/355 30§ ja 34§ ja Kuluttajansuojalaki 1978/38 5. luku 16§ ja 18§). Vahinko voi olla välitön tai välillinen. Välittömiä vahinkoja ovat muun muassa matkakulut, postikulut sekä korjauskulut[138]. Välilliset vahingot sen sijaan on määritelty kauppalaissa ja kuluttajansuojalaissa eri tavoin. Tällaisia kauppalain tarkoittamia vahinkoja voivat olla muun muassa tuotannon tai liikevaihdon vähentyminen tai keskeytyminen tai saamatta jäänyt voitto (Kauppalaki 67§). Kuluttajansuojalain mukaisia välillisiä vahinkoja ovat muun muassa tulonmenetys ja käyttöhyödyn menetys,

[137] Kallioniemi, 2022, s.140-159.
[138] Kilpailu- ja kuluttajavirasto. Vahingonkorvaus tavaran virheestä ja Kilpailu- ja kuluttajavirasto. Vahingonkorvaus palvelun ja virheestä ja viivästyksestä.

lisäksi kuluttaja voi hakea korvausta omaisuudelle aiheutuneista vahingoista (Kuluttajansuojalaki 5. luku 10§ 3 mom. ja 21§). Virheestä syntyvä vahingonkorvaus edellyttää vahingon syntymisen näyttöä.[139] Tekoälyn aiheuttamana vahingon kärsineellä henkilöllä on sama suojan taso kuin olisi kyseessä, jos vahinko olisi syntynyt mistä tahansa muusta teknologiasta[140].

3.3 Käyttäjän ja tarjoajan kantama vastuu

Sopimusperusteisen vastuun sekä kuluttajansuojan ja kauppalain tuomat ehdot näyttävät poistavan vastuun itse tekoälyltä itsenäisenä oikeudellisena yksikkönä. Vastuuta siirretään paljon tekoälyn toimittamaan osalliseen sekä käyttäjälle siltä osin, kun vahingot johtuisivat käyttäjästä itsestään. Tavallinen tilanne siis on, että tarjoajan sopimuskumppani vaatisi tekoälyn tarjoajalta vahingonkorvausta järjestelmän suorittaman toiminnallisuuden takia[141]. Lisätäkseen kansalaisen suojaa tekoälyä koskevissa vahingonkorvausasioissa, Euroopan unioni on antanut direktiiviehdotuksen, ”direktiivi tekoälyyn liittyvästä vastuusta”[142], tekoälyä koskevasta vastuuvelvollisuudesta sopimuksenulkoisissa tilanteissa. Lisäsuojan perusteena on tekoälyn läpinäkymättömyys, autonomia ja kompleksisuus, jotka vaikeuttavat virheen ja vastuussa olevan henkilön tunnistamista[143]. Tavallisessa tilanteessa uhrin tulisi todistaa tekijän suorittaneen väärä toiminto ja että vahinko tapahtui. Virheen ja syy-yhteyden todistaminen voi kuitenkin osoittautua hankalaksi, minkä takia pyritään mahdollistamaan, että uhri voi saada samanlaisen suojan kuin ilman tekoälyä varustettujen tuotteiden ja palvelun vahinkotilanteissa. Lisäksi nykyisten

[139] Kilpailu- ja kuluttajavirasto. Vahingonkorvaus tavaran virheestä ja Kilpailu- ja kuluttajavirasto. Vahingonkorvaus palvelun ja virheestä ja viivästyksestä.
[140] Euroopan komissio, 2020, s.15.
[141] Kallioniemi, 2022, s.167.
[142] Ehdotus Euroopan parlamentin ja neuvoston direktiivi sopimuksenulkoista siviilioikeudellista vastuuta koskevien sääntöjen mukauttamisesta tekoälyyn (direktiivi tekoälyyn liittyvästä vastuusta) COM(2022) 496 final.
[143] Euroopan komissio, 2022, s.1.

säännösten nähdään aiheuttavan oikeudellista epävarmuutta siitä, kuinka näitä tulisi ja tullaan soveltamaan tekoälytapauksissa kansallisesti ja unionin tasolla.[144]

Tarjoajan vastuusta on otettu kantaa myös tekoälysäädösehdotuksessa viitaten kuitenkin nimenomaan suuririskisiin tekoälyjärjestelmiin:

> *”On asianmukaista, että tietty tarjoajaksi määritelty luonnollinen tai oikeushenkilö ottaa vastuun suuririskisen tekoälyjärjestelmän markkinoille saattamisesta tai käyttöönotosta riippumatta siitä, onko kyseinen luonnollinen henkilö tai oikeushenkilö järjestelmän suunnittelija tai kehittäjä”.*[145]

Ehdotuksesta käy ilmi, että tarjoajan olisi ainakin luotava laadunhallinta-, riskinhallinta- ja seurantajärjestelmä, varmistettava arviointimenetelmien suorittaminen sekä laadittava dokumentaatio asiaa koskien[146]. Dokumentaation tulee sisältää tietoja muun muassa tekoälyn kehittämiseen käytetyistä menetelmistä, järjestelmään käytetyt ja integroidut kolmannen osapuolen tarjoamista ennalta koulutetuista järjestelmistä tai välineistä, rakennespesifikaatiot, järjestelmän ja algoritmien yleinen logiikka, periaatteet ja tehdyt oletukset, luokitusvalinnat, järjestelmäarkkitehtuurin kuvaus, koulutusmenetelmät ja -tekniikat, datajoukkojen alkuperä ja laajuus sekä validointimenetelmät.[147]

Viranomainen, joka ottaa käyttöön suuririskisiä tekoälyjärjestelmiä voi hyväksyä ja panna täytäntöön laadunhallintajärjestelmää koskevat säännöt.[148] Ylipäätänsä ihmisen suorittama valvonta nähdään eräänä tärkeimmistä tavoista varmistaa, ettei tekoäly aiheuttaisi vahingollisia vaikutuksia. Näin koska eettinen ja ihmiskeskeinen tekoäly pystyttäisiin luomaan vain ihmisen väliintulolla. Tällaisen ihmisen suorittaman valvonnan luonne ja laajuus riippuisi kuitenkin tekoälyjärjestelmän käyttötarkoituksesta ja vaikutusmahdollisuuksista muun muassa kansalaisiin. Tapoja suorittaa valvontaa voi olla esimerkiksi se, ettei tekoälyn luomat päätökset ole lainvoimaisia ennen kuin ihminen on kyseisen

---

[144] Euroopan komissio, 2022, s.1-3.
[145] Euroopan komissio, 2021a, johdanto-osan kappale 53.
[146] Euroopan komissio, 2021a, johdanto-osan kappale 54 ja artikla 9.
[147] Euroopan komissio, 2021a, liite IV.
[148] Euroopan komissio, 2021a, johdanto-osan kappale 54.

päätöksen validoinut tai ihmisen mahdollisuus minä tekoälyn käyttöhetkenä tahansa astua väliin ja kytkeä tekoälyjärjestelmä pois päältä.[149]

Suuririskisten tekoälyjärjestelmien kohdalla olisi huomioitava käyttäjäryhmän *odotettu* tekninen osaaminen ja ympäristö, jossa järjestelmää olisi tarkoitus käyttää. [150] Voidaan esimerkiksi olettaa, että julkisesti saatavilla oleva kansalaiskäyttöön tarkoitetun järjestelmän käyttäjät eivät välttämättä omaa teknisiä taitoja verraten esimerkiksi virkakäyttöön tarkoitettu järjestelmä, jossa käyttäjät tuntevat tehtävät toimenpiteet ja saavat myös teknistä koulutusta nimenomaisen järjestelmän käytölle. Käyttäjäryhmän osaamistason huomioiminen olisi osa riskien mitigoimista.

Myös käyttäjälle itselleen asetettaisiin vastuuta tekoälyjärjestelmän mahdollisista haittavaikutuksista. Järjestelmää olisi käytettävä käyttöohjeiden mukaisesti, mikä tarkoittaa, että tekoälyn tarjoajalle syntyy velvoite asettaa ohjeet ja muu tarpeellinen dokumentaatio käyttäjän saataville. Samoin olisi ilmoitettava, että käyttäjä on vuorovaikutuksessa tekoälyllä varustetun järjestelmän kanssa, ellei tämä ole jo ilmeistä olosuhteiden ja käyttöyhteyden perusteella. Jos taas tekoälyjärjestelmän yhteydessä käyttäjä altistuu tunteentunnistusjärjestelmälle tai biometriselle luokitusjärjestelmälle tai jos järjestelmä tuottaa tai manipuloi kuva-, ääni- tai videosisältöä, joka muistuttaa selkeäsi olemassa olevia henkilöitä, paikkoja tai tapahtumia ja, joka voi vaikuttaa aidolta, on tämäkin tieto tarjottava käyttäjälle.[151]

Osa tekoälyä koskevaa saataville laitettavaa dokumentaatiota pitäisi olla tiedot perusoikeuksiin kohdistuvista riskeistä, syrjintäriskeistä sekä yleisesti tekoälyn rajoitteita ja toimintakyvykkyyksiä koskevat tiedot. Erityisesti on tiedotettava, mihin tarkoitukseen

[149] Euroopan komissio, 2020, s.21.
[150] Euroopan komissio, 2021a, art 9 alakohta 4.
[151] Euroopan komissio, 2021a, johdanto-osan kappale 70 ja art 29; Tekoälyä käsittelevä korkean tason asiantuntijaryhmä, 2019a, s.12 ja Euroopan komissio, 2020, s.20.

järjestelmä on tarkoitettu ja missä olosuhteissa järjestelmän voidaan olettaa toimivan kuten on tarkoitettu ja linjassa odotetun virheettömyyden tason kanssa.[152]

3.4 Euroopan unionin rekisteri tekoälyjärjestelmille

Euroopan unionissa on tarkoitus perustaa hallintojärjestelmä koko unionin ja kansalliselle tasolle. Unionin tasolla tämä tarkoittaisi Euroopan tekoälyneuvoston perustamista. Neuvosto edistäisi kansallisten valvontaviranomaisten ja komission välistä yhteistyötä sekä antaisi neuvoja ja asiantuntemusta komissiolle tekoälysäädöksen täytäntöönpanoa edistääkseen. Kansallisella tasolla kukin jäsenvaltio perustaisi tai nimeäisi toimivaltaiset viranomaiset tekoälysäädöksen täytäntöönpanon varmistamiseksi. Kansallinen viranomainen antaisi ohjeita ja neuvoja tekoälyjärjestelmää koskien.[153]

Tekoälyjärjestelmien tarjoajille on pyrkimys luoda seuranta- ja raportointivelvoitteet järjestelmien markkinoille saattamisen jälkeistä seurantaa ja raportointia sekä tekoälyyn liittyviä vaaratilanteiden ja toimintahäiriöiden tutkintaa varten.[154] Vaaratilanteista tai toimintahäiriöistä ilmoitettaisiin kansalliselle viranomaiselle heti kun tekoälyn tarjoaja on saanut tällaisesta tiedon. Ilmoituksen jälkeen kansalliselle viranomaiselle syntyisi velvoite tutkia vaaratilanne tai häiriö, kerätä tarpeelliset tiedot ja toimittaa ne komissiolle.[155] Lisäksi suuririskisten tekoälyjärjestelmien tarjoajia olisi vaadittava rekisteröimään tällainen tekoälyjärjestelmä Euroopan unionin tietokantaan suuririskisistä tekoälyjärjestelmistä. Tietokanta olisi komission perustama ja hallinnoima sekä komissio toimisi tietokannan rekisterinpitäjänä.[156] Rekisteröintiä varten toteutetaan erillinen järjestelmä, jonka kautta toimivaltaiset viranomaiset, käyttäjät ja muut henkilöt voivat halutessaan tarkistaa, täyttääkö suuririskinen tekoälyjärjestelmä tekoälysäädösehdotuksessa esitetyt vaatimukset.[157]

---

[152] Euroopan komissio, 2021a, johdanto-osan kappale 47 ja Euroopan komissio, 2020, s.20.
[153] Euroopan komissio, 2021a, art 57-59.
[154] Euroopan komissio, 2021a, s.16
[155] Euroopan komissio, 2021a, s.13
[156] Euroopan komissio, 2021a, art 60.
[157] Euroopan komissio, 2021a, s 11-12.

4 Tekoäly ja Data

Datan ja erityisesti digitaalisessa muodossa olevan datan merkitys on kasvanut nyky-yhteiskunnassa uudelle tasolle, kun halu tehostaa toimintoja automaatioteknologialla on lisääntynyt. Data on perusedellytys tekoäly automaatiolle sekä tekoälyn eräs tärkeimmistä rakennusosista.[158] Ilman dataa ei ole tekoälyä[159]. Kuten jo tekoälyn määritelmää avatessani kävi ilmi, tekoäly on oppiva kokonaisuus. Oppiminen tapahtuu dataa analysoimalla ja käsittelemällä. Esimerkiksi koneoppimista hyödyntävä järjestelmä tunnistaa säännösmukaisuuksia tarjotusta tietojoukosta eli datasta ja soveltaa näitä säännönkaltaisia havaintoja uuteen tietojoukkoon[160].

Eräs soveltamistapa voisi olla kuvantunnistuksen toiminto eli konenäkö, jonka avulla algoritmi luokittelee esineitä. Tässä tapauksessa algoritmille syötetään esimerkkidatana jo valmiiksi luokiteltuja kuvia ja näistä löytyvien ”sääntöjen” avulla tekoäly pystyy luokittelemaan tälle kyseiselle tekoälylle täysin uusina näkyviä kuvia.[161] Konenäköä pystyttäisiin hyödyntämään muun muassa lääketieteessä, psykiatriassa tai rikostutkinnassa esimerkiksi tunnistamaan ihmisen tunnetiloja tai jopa geneettisiä sairauksia kasvojenpiirteistä sekunneissa sen sijaan, että samaan diagnoosiin menisi ihmiseltä jopa 16 vuotta. Finnair on esimerkiksi vuonna 2017 järjestänyt matkustajilleen vapaehtoisen kasvojentunnistuskokeilun[162]. Kokeilussa vapaaehtoiset latasivat kännykkäänsä sovelluksen ja ottivat itsestään kuvan. Kun samainen asiakas saapui kentälle, lentoaseman kameran piti tunnistaa kyseinen henkilö jo jonossa ja päästää suoraan tarkastusportin läpi.[163]

Vaatimuksena tällaiselle ihmisen ”lukemiselle” on, että algoritmi opetetaan rekisteröimään muun muassa kasvojen perus- ja mikroilmeitä tai sydämen sykkeen aiheuttamia

[158] Kaarlejärvi ja salminen, 2018, s.30 ja Euroopan komissio, 2020, s.19.
[159] Euroopan komissio, 2020, s.19.
[160] Euroopan komissio, 2018, s.10.
[161] Euroopan komissio, 2018, s.11.
[162] Siukonen ja Neittaanmäki, s.75.
[163] Siukonen ja Neittaanmäki, s.75.

värivaihteluita[164]. Ihmisestä itsestään saatavaa dataa voidaan kerätä puheesta (voimakkuus, sanavalinnat…), äänestä (intonaatio, äänensävy…), kasvoista (hymyt, silmänliikkeet…), sensoreista (galvaaninen ihoreaktio, lämpötila, verenpaine, hengitystiheys…) tai lisälaitteiden (reaktioaika, kognitio, liikkeet…) avulla ja muuntamaan se digitaalisesti käytettäväksi dataksi. Näin on tehty esimerkiksi etelä-Kalifornian yliopistossa, kun sovellus kehitettiin käyttämään 74 eri akustista ominaisuutta ennustamaan eripuraa parisuhteessa. Sovellus onnistui ennustuksissaan jopa paremmin kuin terapeutit itse. Edellä esitetyn kaltaista heikkoa tekoälyä on ollut jo laajahkosti käytössä jo arkielämässä muun muassa eri kasvojentunnistusvälineenä avaamaan älypuhelin ilman kirjoitettua salasanaa.[165]

4.1 GDPR ja kyberturvallisuusasetus

Tietosuojalaki (2018/1050) on Suomen omaa kansallista lainsäädäntöä, jonka tarkoituksena on lain ensimmäisen pykälän mukaisesti täsmentää ja täydentää Euroopan parlamentin ja neuvoston antamaa yleistä tietosuoja-asetusta (2016/679) luonnollisten henkilöiden henkilötietojen suojasta henkilötietoja käsiteltäessä. Yleisen tietosuoja-asetuksen 74 kohdan mukaan on vahvistettava rekisterinpitäjän vastuu suorittamastaan tai puolestaan suoritetusta henkilötietojen käsittelystä. Kohdan mukaisesti rekisterinpitäjän tulisi erityisesti olla velvollinen toteuttamaan tehokkaat ja asianmukaiset toimenpiteet, joita toteutettaessa olisi huomioitava rekisteröidyn oikeuksiin ja vapauksiin kohdistuvan riskin todennäköisyys ja vakavuus. Toimenpiteet määritellään objektiivisesti tietojenkäsittelyn luonteen, laajuuden, asiayhteyden ja tarkoituksen mukaan.

Koska tietojärjestelmät nimensä mukaisesti merkitsevät tiedon, usein henkilötiedon käsittelyä, syntyy tässä linkki tietojärjestelmien vaatimuksista eli, mitä tulee huomioida tällaisia järjestelmiä suunniteltaessa samoin kuin sen ylläpidossa (Yleinen tietosuoja-asetus,

[164] Siukonen ja Neittaanmäki, 2019, s.75 ja Topol, 2019, s.57.
[165] Topol, 2019, s.82 ja 168.

art. 24 kohta 1 ja art. 32 kohta 1 alakohta d). Eräinä asetuksen tarkoittamina toimenpiteinä voidaan pitää muun muassa henkilötietojen käsittelyn minimointia, pseudonymisointia ja salausta sekä kykyä palauttaa tietojen saatavuus. (Yleinen tietosuoja-asetus, johdanto-osan kappale 78 ja art. 32). Oikeuksiin ja vapauksiin kohdistuvina riskeinä voidaan sen sijaan pitää tietojen tuhoutumista, häviämistä, muuttamista sekä luvatonta luovuttamista tai pääsyä. Edellä kuvattuja riskejä aiheuttaa esimerkiksi tietojen siirto, tallennus sekä muu käsittely yleisesti. (Yleinen tietosuoja-asetus, johdanto-osan kappale 83).

Ottaen huomioon ”siirto”, ”tallennus” ja ylipäätänsä ”käsittely”, joka voi asetuksen 4. artiklan 1. momentin 2. kohdan mukaan tarkoittaa myös muun muassa tiedon hakemista, kyselyä tai muuttamista, on tietojärjestelmän kokonaisuudessaan huomioitava kaikki se, mikä voisi loukata oikeuksia ja vapauksia, kuten henkilötietojen suojaa tässä tapauksessa. Tietojärjestelmän suunnitteluun liittyy siis hyvin laaja kirjo huomioitavia asioita, ei ainoastaan tietosuoja vaan esimerkiksi myös oikeus tarkastella omia tietojaan. Lisäksi tietosuoja ei kata vain käskyä, että tietoja ei saa levitä ulkopuolisen käsiin, vaan tietojärjestelmän tulee toimia preventiivisesti ja näin ollen sen suunnittelijan tulee aktiivisesti pyrkiä vähentämään tietosuojariskejä esimerkiksi minimoimalla käsiteltävän tiedon määrää tai luoda mahdollisimman vaikeasti ohitettava varmennuskeino tietokantaan päästessä käsiksi.

Muuta tietosuojan kannalta olennaista kansainvälistä lainsäädäntöä on esimerkiksi kyberturvallisuusasetus (2019/881). Tämän asetuksen mukaan on asetettu oletusarvoinen tietoturva. Käytännössä tietoturvan oletusarvo tarkoittaa, että toimijoiden on konfiguroitava suunnittelemansa tieto- ja viestintätekniikan tuotteet, palvelut ja prosessit niin, että käyttäjälle ei jää tehtäväkseen muokata noita asetuksia asianmukaiselle tietoturvan tasolle. Suunniteltujen hyödykkeiden ja prosessien ei kuitenkaan vaadita olevan täydellisesti tietosuojan takaavia, vaan asetuksen johdanto-osan kappaleessa 13 mainitaan, että oletusarvoinen turvallisuus ei saisi edellyttää mittavaa konfigurointia eikä teknistä

erityisosaamista tai muuta kuin ilmeistä käyttäytymistä käyttäjältä. Näin ollen käyttäjälle jää siis omakohtaista vastuuta tietoturvallisuuden saavuttamisesta.

Kansallisella tasolla muun muassa laki julkisen hallinnon tiedonhallinnasta asettaa järjestelmille tietoturvavaatimuksia. Käytännössä tietojärjestelmän tulisi olla tietoturvallinen koko sen elinkaaren ajan aina suunnittelusta ylläpitoon asti (Laki julkisen hallinnon tiedonhallinnasta 906/2019 (Tiedonhallintalaki) 13§). Tietoturvallisuus tässä laissa tarkoittaa, että varmistetaan tietoaineiston muuttumattomuus, alkuperäisyys, ajantasaisuus, virheettömyys, saatavuus ja käyttökelpoisuus; tietoaineisto suojataan teknisiltä ja fyysisiltä vahingoilta; ja tietoaineistojen saatavuutta rajoitetaan vain, jos tiedonsaantia tai käsittelyoikeutta on laissa erikseen rajoitettu (Tiedonhallintalaki 15§)

4.2 Tiedonhallintalaki

Tiedonhallintalaki asettaa tietojärjestelmille uusina ominaisuuksina vikasietoisuuden ja asiakirjajulkisuuden, sekä ottaa kantaa tietoaineistojen säilytysaikoihin (Laki julkisen hallinnon tiedonhallinnasta 906/2019 (Tiedonhallintalaki) 13§). Tietoaineistojen säilytyksessä tulee toteuttaa laissa säädettyjä vaatimuksia, tai jos tällaisia ei ole asetettu, tietoaineistoa tulee säilyttää vain sen aikaa, kun on tarkoituksenmukaista. Eli aineistoa tulisi säilyttää vain, mikäli sille on käyttötarkoitus, tai kun kyse on esimerkiksi henkilön etujen, oikeuksien, velvollisuuksien ja oikeusturvan toteuttamisesta tai sopimuksen oikeusvaikutuksesta. (Tiedonhallintalaki 21§). Olennaisena asiana tietojen käsittelyn rajaamisen kannalta on rekisteröidyn oikeuksien lisäksi tietoon kohdistuva tietosuojariski. Riskiä voidaan näet pienentää, kun rajataan tiedon käsittelyä asiointipalvelussa ja kun se rajataan nimenomaan käyttötarkoituksen kannalta välttämättömään tietojoukkoon.[166]

Sikäli kun määritetään säilytysajasta, tarkoittaa se, että lopulta tietoaineisto on tuhottava tai arkistoitava, joten järjestelmästä tulisi voida turvallisesti poistamaan tietoa

[166] Valtiovarainministeriö, 2017, s.27.

tuhoamatta muuta tietoaineistoa samalla. Näin säätää esimerkiksi arkistolaki (1994/831) pykälässä 13 sekä laki hallinnon yhteisistä sähköisen asioinnin tukipalveluista pykälässään 15. Kyseisen tukipalvelulain mukaan käyttäjäorganisaation, eli tietoja käsittelevän, on välittömästi hävitettävä yksilöivää tunnusta koskeva tieto, jos organisaatiolla ei ole oikeutta käsitellä tuota tietoa.

Alkujaan tarkoituksenmukaisuusvaatimus ja säilytysajan minimaalisuus on esitetty yleisessä tietosuoja-asetuksessa: rekisterinpitäjän on asetettava määräajat henkilötietojen poistoa tai niiden säilyttämisen tarpeellisuuden määräaikaistarkastelua varten. Tämän tarkoitus on varmistaa, ettei henkilötietoja säilytetä pidempään kuin on tarpeen. Lisäksi henkilötietoja voidaan käsitellä vain, jos käsittelyn tarkoitusta ei voida toteuttaa muilla keinoin. (Yleinen tietosuoja-asetus, art. 17 ja johdanto-osan kappaleet 39 ja 65).

Tietoaineiston saatavuuden osalta tiedon on oltava helposti saatavilla ja ymmärrettävissä ja tieto tulee voida antaa sähköisessä muodossa (Yleinen tietosuoja-asetus, johdanto-osan kappale 58). Yleisen tietosuoja-asetuksen tarkoituksena on säätää mekanismeista, joilla rekisteröity voi pyytää pääsyä henkilötietoihinsa ja esittää henkilötietojensa oikaisu- tai poistamisvaatimuksen. Rekisterinpitäjälle lankeaa tätä myötä velvoite tarjota keino esittää näitä edellä mainittuja pyyntöjä, mutta ennen kaikkea velvoite myöskin vastata tuohon pyyntöön viimeistään kuuden kuukauden kuluessa. Mitä tulee oikeuteen päästä käsiksi henkilötietoihinsa, tulisi rekisteröidyllä olla mahdollisuus päästä etäyhteyden kautta suojatussa järjestelmässä käsiksi omiin tietoihinsa (Yleinen tietosuoja-asetus, johdanto-osan kappaleet 59 ja 63).

4.3 Millaista dataa voidaan käyttää tekoälyn muovaamiseen?

Tekoälyn opettamiseen käytettävä data ei saisi olla puutteellista, vanhentunutta, keksittyä tai saastunutta eikä sitä saisi olla liian vähäisissä määrin[167]. Yksinkertaisuudessaan

[167] Siukonen ja Neittaanmäki, 2019, s.51-54 ja Virtanen ja Stenvall, 2014, s 55.

tekoälyalgoritmi, joka on esimerkiksi opetettu tunnistamaan vain punaisia tomaatteja kuvista, ei pysty tunnistamaan keltaista tomaattia tomaatiksi ollenkaan. Mikäli tekoäly olisi kyseisessä tapauksessa opetettu laajemmalla datajoukolla erilaisista tomaateista, voisi algoritmi tunnistaa myös erivärisiä tomaatteja[168]. Tämän tyyppinen tilanne on muodostunut muun muassa Googlen kuvantunnistuksella, joka kategorisoi kaksi tummaihoista miestä gorilloiksi[169].

Tekoälyalgoritmin vuorovaikutus, samoin kuin sen tekemät päätökset ja toimenpiteet, riippuvat algoritmin vastaanottamasta datasta jo oppimisvaiheessa[170]: mitä suurempi tietojoukko algoritmilla on käytössään, sitä hienovaraisempia yhteyksiä tekoäly voi löytää[171]; samoin, mitä monipuolisempaa dataa tekoäly saa oppiessaan, sitä täsmällisempiä tuotoksia se voi antaa ulos[172]. Mikäli tekoälyjärjestelmää ei olla koulutettu laadukkaalla datalla, se ei voi täyttää riittäviä vaatimuksia tarkkuutensa ja luotettavuutensa suhteen eikä voida myöskään luottaa, että järjestelmä olisi suunniteltu ja testattu asianmukaisesti ennen markkinoille saattamista tai käyttöönottoa. Tällainen järjestelmä on omiaan aiheuttamaan perus- ja ihmisoikeusvaikutuksia esimerkiksi valikoimalla henkilöitä syrjivästi tai muutoin epäoikeudenmukaisesti.[173]

Datan laadun ja määrän voidaan siis todeta olevan ratkaisevia tekijöitä tekoälyn toiminnan kannalta[174] ja siksi myös Euroopan unionissa on tunnistettu tarve varmistaa tekoälyn koulutuksessa käytettävän datan vaatimukset yhdessä Euroopan unionin arvojen ja sääntöjen kanssa[175]. Euroopan komission ehdottamassa tekoälysäädöksessä pyydetään otettavan huomioon koulutus-, validointi- ja testausjoukoissa niiden käyttötarkoituksen edellyttämässä laajuudessa ominaisuudet ja osatekijät suhteessa tekoälyn käyttökohteen maantieteelliseen, käyttäytymiseen liittyvään tai toiminnalliseen ympäristöön tai

[168] De Bruyne ja Vanleenhove, 2021, s.10.
[169] De Bruyne ja Vanleenhove, 2021, s.11.
[170] Euroopan komissio, 2020, s.19 ja Euroopan komissio, 2018, s.10.
[171] Euroopan komissio, 2018, s.10.
[172] Siukonen ja Neittaanmäki, 2019, s.51-54.
[173] Euroopan komissio, 2021a, johdanto-osan kappale 38.
[174] Siukonen ja Neittaanmäki, 2019, s.51-54 ja COM 2021 205, s.5
[175] Euroopan komissio, 2020, s.19 ja Euroopan komissio, 2021a, johdanto-osan kappale 38.

asiayhteyteen. Lisäksi datajoukkojen olisi oltava merkityksellisiä, edustavia ja virheettömiä suhteutettuna käyttötarkoitukseen.[176]

Komission ehdotuksella täydennetään myös syrjimättömyyttä koskevaa voimassa olevaa Euroopan unionin lainsäädäntöä erityisvaatimuksilla, joilla minimoitaisiin algoritmisen syrjinnän riskiä. Nämä syrjintää ehkäisevät vaatimukset koskevat nimenomaisesti tekoälyjärjestelmän kehittämisessä käytettävien datajoukkojen suunnittelua ja laatua.[177] Datan edellytyksiä koskevista periaatteista huolimatta säädös ei tarkenna, millainen datajoukko voidaan todeta olevan tämän määritellyn laadun mukaista ja täten muun muassa tarpeeksi kattavaa, jotta ne täyttäisivät kaikki eri tilanteet eivätkä johtaisi mahdollisesti datasta johtuvaan diskriminointiin. Komissioon verraten näitä samoja laadukkaan datan ominaisuuksia ovat kuvanneet Kaarlejärvi ja Salminen vuoden 2018 julkaisussaan ”Älykäs taloushallinto: Automaation aika”. Heidän mukaansa laatua voidaan kasvattaa muuttamalla paperilla olevaa dataa digitaaliseen muotoon, samoin ei-rakenteisessa muodossa oleva data rakenteiseen muotoon, korjaamalla datan virheitä ja rikastamalla dataa tarpeellisilla elementeillä, varmistamalla datan oikeellisuus kontrolleilla, täsmäysrutiineilla ja päivityksillä, parantamalla datan oikea-aikaisuutta sekä yhtenäistämällä ja yhdenmukaistamalla dataa[178]. Myös komission tekoälysäädösehdotus antaa muun muassa oikeutuksen rikastaa, puhdistaa ja yhdistellä datajoukkoja[179]. Määritelmät eivät kuitenkaan kuvaa, mikä voidaan nähdä riittävänä varmistuksena laadusta kullekin laadukkaan datan ominaisuudelle.

4.4 Euroopan unionin yhteinen data-avaruus

Euroopan unionin julkaisussa ”Coordinated Plan on Artificial Intelligence (COM2018/795)” luotiin ajatus koko unionin alueen kattavasta data-alueesta.[180] Sikäli

[176] Euroopan komissio, 2021a, johdanto-osan kappaleet 44-45.
[177] Euroopan komissio, 2021a, s.4.
[178] Kaarlejärvi ja Salminen, 2018, s. 68-69.
[179] Euroopan komissio 2021a, art 10.
[180] Euroopan komissio, 2021b, s.11.

kun data on tekoälyn kehityksen polttoaine, pyrkimys on lisätä datan saatavuutta ja laatua tasaisesti koko Euroopan unionin alueella[181]. Näin mahdollistetaan digitaalisten palveluiden saavutettavuus tarjoamalla hyvälaatuista tietoa silloin kun sitä tarvitaan tiettyjen lainsäädännöllisten puitteiden valossa[182]. Tällaisen avaruudellisen tilan luomisen seurauksena tieto ja data kulkisivat rajat ylittävällä tavalla helposti ja täysin laillisesti.

Edellä mainitussa komission julkaisussa tunnistettiin ongelmakohdaksi datan eristyminen yksittäisille tekijöille sen seurauksena, että yritykset eivät jaa dataansa keskenään[183]. Ilmiö on tosin luonnollinen ottaen huomioon yksityisellä sektorilla vallitsevan kilpailutilanteen ja datan tärkeyden resurssina itsenään. Tätä ajatellen unioni loisi insentiivin datan jakamiselle sekä selkeät ja diskriminoimattomat säännöt dataan pääsemiksi ja sen käytölle kuitenkin niin, että säännöt noudattavat unionin arvoja ja säädäntöä[184]. Data-avaruuden tärkeys perusteltiin myöhemmin vielä tekoälysäädösehdotuksessa sanoen tämän tarjoavan luotettavan, vastuullisen ja syrjimättömän pääsyn korkealaatuiseen dataan[185].

Datan jakamisen helpottamiseksi ja data-avaruuden luomiseksi Euroopan unionin sisällä päätettiin luoda datalle omat sisämarkkinat[186]. Datan sallitaan täten liikkua vapaasti rajojen yli ja sektoreiden välillä eli myös julkiselta yksityiselle ja toisin päin riippumatta sen fyysisestä säilytyspaikasta[187]. Sama olisi myöhemmin sallittavan myös kolmansiin maihin tiettyjen, muun muassa turvallisuutta koskevien, ehtojen täyttyessä. Yhdenmukaiset puitteet datan jakamiselle tulisivat osana sisämarkkinoita ja täten voitaisiin ohjata datatalouden tilaa pisteeseen, jossa yritykset kilpailevat datan laadulla määrän sijaan. Tähän mennessä julkisissa tietokannoissa olevaa dataa, esimerkiksi luottamuksellista dataa,

---

[181] Työ- ja elinkeinoministeriö, 2017, 42.
[182] Työ- ja elinkeinoministeriö, 2017, 53.
[183] Euroopan komissio, 2021b, s.12.
[184] Euroopan komissio, 2021b, s.12 ja Euroopan komissio, European data strategy.
[185] Euroopan komissio, 2021a, johdanto-osan kappale 45.
[186] Euroopan komissio, 2021b, s12 ja Euroopan komissio, European data strategy.
[187] Euroopan komissio, 2021b, s12; Euroopan komissio, European data strategy; Euroopan parlamentin ja neuvoston asetus (EU) 2022/868 eurooppalaisen datan hallinnoinnista ja asetuksen (EU) 2018/1724 muuttamisesta(datanhallinta-asetus) johdanto-osan kappale 1.

salassa pidettävää tilastotietoa, teollis- ja tekijänoikeuksilla suojattua dataa sekä liikesalaisuuksia tai henkilötietoa, ei olla asetettu saataville, vaikka se olisi ollut mahdollista. Komission mukaan tarve on tästä syystä luoda edellytykset datan jakamiselle ja käytölle. (Datanhallinta-asetus (EU) 2022/868, johdanto-osan kappaleet 2,3 ja 6).

Datan liikkuvuutta vielä lisätäkseen ja osana Euroopan datastrategiaa komissio ehdotti uudenlaista eurooppalaista hallintotapaa datan jakamiselle eri toimialojen ja EU-maiden välillä[188]. Kyseessä on EU:n datanhallinta-asetus, joka on annettu toukokuussa 2022 ja on tällä hetkellä voimassa oleva asetus. Datanhallinta-asetus käsittelee datan uudelleenkäyttöä niin kansallisessa kuin kansainvälisessä merkityksessä sekä esittelee uudelleenkäyttöön liittyvät keskeisimmät toimijat ja heidän velvollisuutensa. Datan uudelleenkäyttö pitää sisällään datan käytön esimerkiksi kaupalliseen tarkoitukseen poiketen datan alkuperäisestä käyttötarkoituksesta julkisessa tehtävässä. Data on näissä tilanteissa siis julkisen sektorin elimen hallussa ja käyttäjänä toimii luonnollinen henkilö tai oikeushenkilö. (Datanhallinta-asetus (EU) 2022/868, art. 2 kohta 2)

Erityisenä huomiona säädös mahdollistaa organisaation rekisteröinnin data-altruismipohjaiseksi, kun kyseessä on oikeushenkilö, joka tukee yleisen edun mukaisia tavoitteita asettamalla laajamittaisesti dataa saataville data-altruismi pohjanaan (Datanhallinta-asetus (EU) 2022/868 johdanto-osan kappale 3 ja art. 18). Näitä organisaatioita nimitetään ”Unionissa tunnustetuiksi data-altruismipohjaisiksi organisaatioiksi” ja tällaiseksi rekisteröinti olisi voimassa koko unionin alueella (Datanhallinta-asetus (EU) 2022/868 johdanto-osan kappaleet 3 ja 46).

Tekoälysäädösehdotus lausumissaan toteaa pääsyn olevan tarkoitettua tekoälyjärjestelmien koulutusta, validointia ja testausta varten sulkien siis pois useita datan käyttötarkoituksia. Rajaus ei kuitenkaan tarkoita, että tähän tarkoitukseen käytettävä data ei voisi olla autenttista henkilötietoa tai yksilöön liitettävää tietoa.[189] Komissio ehdottaa

[188] Euroopan komissio, European data strategy.
[189] Euroopan komissio, 2021a, johdanto-osan kappale 45.

käytännössä toimialakohtaisia yhteisiä eurooppalaisia data-avaruuksia, jotka mahdollistavat datan jakamisen samalla yhdistäen samankaltaisen datan samaan kokonaisuuteen. Erityisinä vaatimuksina data-avaruuksille ovat datan löydettävyys, saavutettavuus, yhteen toimivuus ja uudelleenkäytettävyys sekä tietenkin kyberturvallisuuden korkea taso.[190] Esimerkkinä ehdotus antaa terveydenhuollon alan datan, joka kuuluisikin ”terveysdata-avaruuteen”. Terveysdata olisi täten syrjimättömästi saatavilla ja tekoälyalgoritmien koulutusta varten niin, että datan jaon aikana huomioidaan yksityisyyden suoja, turvallisuus, oikea-aikaisuus, läpinäkyvyys ja luotettavuus.[191]

Esimerkki viittaa, että tekoälyn kehittämiseen edellytetty data on oikeita henkilöitä koskevaa tietoa. Yksilöille annettaisiin kuitenkin tässäkin tilanteessa oikeus, ja tämän oikeuden toteutumista varten tarvittavat työkalut ja osaaminen, säilyttää täysi hallinta heitä koskevasta datasta[192]. Julkisen sektorin hallussa olevien tiettyjen datajoukkojen käsittelyn olisi tapahduttava rekisteröidyn oikeuksia kunnioittaen ja julkisen sektorin olisi varmistettava, että sekä luonnollisten henkilöiden että oikeushenkilöiden oikeudet suojataan täysin. Erityisesti henkilötiedot, kaupallisesti arkaluonteiset tiedot ja teollis- ja tekijänoikeuksiin liittyvä tieto suojataan. Sama pätee, kun tietoja siirretään kolmansiin maihin. Kolmansiin maihin tietoja voidaan kuitenkin siirtää vain, mikäli kyseisessä maassa varmistetaan vastaava tietosuojan taso kuin unionin oikeudessa vaaditaan. (Datanhallinta-asetus (EU) 2022/868, johdanto-osan kappaleet 19-22). Lisäksi näihin uudelleenkäytettäväksi annettuihin tietoihin sovelletaan edelleen yleisen tietosuoja-asetuksen esittämiä vaateita henkilötiedoille ja niiden käsittelylle (Datanhallinta-asetus (EU) 2022/868, johdanto-osan kappaleet 4 ja 35).

Tekoälyä koskevassa valkoisessa kirjassa on tunnistettu yleisesti ongelma tekoälyn mahdollisesta kyvykkyydestä luoda linkkejä ja yhdistellä tietoa sekä deanonymisoida eli uudelleentunnistaa anonymisoitua dataa täten tunnistaen henkilön, jota tieto koskee. Sikäli

[190] Euroopan parlamentin ja neuvoston asetus (EU) 2022/868 eurooppalaisen datan hallinnoinnista ja asetuksen (EU) 2018/1724 muuttamisesta (datanhallinta-asetus) johdanto-osan kappale 2.
[191] Euroopan komissio, 2021a, johdanto-osan kappale 45.
[192] Euroopan komissio, European data strategy.

jos tekoäly kykenee itsenäisesti yhdistelemään tietoa, luo tämä riskin henkilöön kohdistuvasta laajemman tietoprofiilin luomisesta kuin mitä yksittäisen datajoukon avulla pystyttäisiin yhdistämään henkilöön.[193] Tekoäly saattaa siis aiheuttaa tietosuojariskin sellaisenkin datajoukon kohdalla, joka ei sisällä henkilötietoja.

4.5 Tietojen käsittelyn periaatteet

Yleinen tietosuoja-asetus asettaa kuusi tietojen käsittelyn periaatteetta. Nuo periaatteet ovat kirjattuina GDPR:n viidenteen artiklaan tiivistetysti: lainmukaisuus, kohtuullisuus ja läpinäkyvyys; käyttötarkoitussidonnaisuus; tietojen minimointi; täsmällisyys; säilytyksen rajoittaminen; eheys ja luottamuksellisuus. Lisäksi rekisterinpitäjällä on osoitusvelvollisuus, että edellä mainittuja periaatteita noudatetaan käsitellessä henkilötietoja.

Käsittelyn lainmukaisuudella viitataan tietojen käsittelyn oikeutukseen. Rekisterinpitäjällä on oikeus käsitellä tietoja käytännössä vain, kun jokin seuraavista ehdoista täyttyy: rekisteröity antaa tietojen käsittelylle suostumuksensa; käsittelylle on sopimusperusteinen peruste; käsittelyllä pannaan täytäntöön lakisääteisiä vaatimuksia; käsittelyllä suojataan yleistä etua tai luonnollisen henkilön elintärkeitä etuja; tai käsittely toteuttaa oikeutettua etua (GDPR art. 6 kohta 1). Mikäli käsittelyperusteena on suostumus, rekisteröidyn oikeuksiin kuuluu oikeus peruuttaa suostumuksensa eli tällä tavalla estää omien tietojensa käsittely suostumuksen piiriin kuuluvien tietojen osalta (GDPR art. 7 kohta 3).

Käyttötarkoitussidonnaisuus taas asettaa henkilötietojen käsittelylle ehdon, jonka mukaan tietoja voidaan kerätä vain tiettyä, etukäteen nimettyä laillista tarkoitusta varten. Tämän ehdon tulisi olla oletusarvo tietoja käsiteltäessä eli rekisterinpitäjän on varmistettava sen toteutuminen omilla toimenpiteillään. Lisäksi tietojen minimointi kieltää tietojen käytön, ellei se ole tarpeellista suhteessa siihen tarkoitukseen, jota varten tietoja käsitellään. (GDPR art. 1 kohta 1 alakohta b ja c ja art. 25 kohta 2). Edellä mainittujen

[193] Euroopan komissio, 2020, s.11.

kohtien takia, henkilötietoja ei voida käsitellä muuhun kuin ennestään määritettyä tarkoitusta varten muutoin käsittelyn rikkoen Euroopan unionin perusoikeuskirjan 8 artiklaa[194].

Mikäli henkilötietojen säilyttämisen ja keräämisen tarve tuohon ennalta määritettyyn tarkoitukseen ei ole enää tarpeellista, rekisterinpitäjän on omatoimisesti hävitettävä nuo henkilötiedot, joiden käsittelylle ei löydy enää lainmukaista tarkoitusta (GDPR art. 5 kohta 1 alakohta e ja art. 17 kohta 1 alakohta a). Tätä nimitetään säilyttämisen rajoittamisen periaatteeksi. Henkilötietoja käsittelevällä on velvollisuus varmistaa, ja aktiivisesti myötävaikuttaa, että epätarkat tai virheelliset henkilötiedot poistetaan tai oikaistaan pikimmiten (GDPR art. 5 kohta 1 alakohta d). Jotta tietojen täsmällisyyden periaate toteutuu, olennaiseen osaan nousee yleisen tietosuoja-asetuksen artikla 15 rekisteröidyn oikeudesta saada pääsy tietoihin ja saada tietää, mikäli häntä koskevia henkilötietoja käsitellään tai, että niitä ei käsitellä. Pyydettäessä rekisteröidyllä on oikeus saada kaikki käsittelyä koskevat tietonsa tiiviisti, helposti ymmärrettävässä ja saatavilla olevassa muodossa ja yksinkertaisella kielellä (GDPR art. 12 kohta 1). Rekisteröidyn on näet hankala oikaista ja korjata itseään koskevia tietoja muussa tapauksessa. Kyseisen artiklan tärkeys korostuu myös läpinäkyvyyden periaatteen osin rekisteröidyn pystyessä varmistumaan rekisterin sisällöstä ja tietojen käyttötarkoituksen toteutumisesta.

Tietojen oikaisemisesta ja poistamisesta eli oikeudesta tulla unohdetuksi säädetään lisää artikloissa 16 ja 17. Viimeisenä periaatteena eheys ja luottamuksellisuus vaatii, että käsittelyssä tulee varmistaa sen turvallisuus muun muassa suojaamalla tieto luvattomalta ja lainvastaiselta käsittelyltä sekä vahingossa tapahtuvalta häviämiseltä, tuhoamiselta tai vahingoittumiselta. Eheyden ja luottamuksellisuuden turvaamiseksi voidaan hyödyntää organisatorisia tai teknisiä toimia, kuten tietojen pseudonymisointia ja salausta, oikeastaan artiklan 24 mukaan nämä organisatoriset ja tekniset toimenpiteet ovat pakollisia. Rekisterinpitäjän ja henkilötietojen käsittelijän tulisi myös pystyä palauttamaan tiedot saataville sekä luoda menettely, jolla testataan, tutkitaan ja arvioidaan teknisten ja

[194] McDermott, 2017, s.2.

organisatoristen toimenpiteiden tehokkuutta. (GDPR art. 5 kohta 1 alakohta f sekä art. 32 kohta 1). Kyky palauttaa data ei pelkästään toteuta tiedon eheyden ja saatavuuden periaatetta, vaan se myös mahdollistaa rekisteröidyn oikeuden pyytää itseään koskeva tieto.

5 Tietosuoja perusoikeutena

Tietosuoja perustavanlaatuisena oikeutena saa asemansa muun muassa jo Euroopan unionin perusoikeuskirjasta. Samoin vuonna 2016 annettu yleinen tietosuoja-asetus (GDPR 2016/679) on perusluonteeltaan tarkoitettu suojelemaan luonnollisen henkilön henkilötietoja samalla tunnustaen henkilötietojen suojan luonnollisen henkilön oikeutena ensimmäisessä artiklassaan. Suomen omalla kansallisella tasolla tietosuojan asema perusoikeutena tunnustetaan perustuslaissa yksityiselämän suojana.

Perinteisessä mielessä valtion hallinnolla on suuri rooli käytäntöjen ja säädösten luomisessa. Tämä pätee myös kyberturvallisuuden kansallisesta säätämisestä[195]. Sikäli kun kyse on säädöstasoisesta suojan luomisesta, hallinto joutuu vastaamaan kyberavaruuden turvallisuudesta: yhteyden suojaaminen, perusoikeuksien kunnioitus ja luotettavuuden säilyttäminen[196]. Vastuu ei tästä huolimatta ole yksinomaan julkiselle vallalle osoitettu, kun huomioidaan, että korkeatasoinen tietosuoja voidaan saavuttaa vain, jos arvoketjun jokainen palanen, esimerkiksi laitevalmistajat ja ohjelmistokehittäjät, priorisoivat tietosuojan korkealle[197].

Tietoturvatoimenpiteiden olisi siis tarkoituksenmukaista olla tarkkaan harkittuja julkisen sektorin lisäksi myös yksityisen sektorin toimijoiden toimesta eritoten, kun kyseessä on IT- sektori, joka tuottaa ja ylläpitää tietojärjestelmiä ja laitteita[198]. Edelleenkään vastuu tietosuojasta ei pysähdy tähän kohtaan, vaan se jatkuu yhä loppukäyttäjien huoleksi. Loppukäyttäjän käsitellessä dataa, on hän myöskin potentiaalinen uhka tietosuojalle, joten heidänkin tulee olla tietoisia tietoverkon riskeistä ja omalta osaltaan suojata tietoa annettujen tietosuojaohjeiden mukaisesti.

[195] Rossini ja Green, 2015, s.2
[196] European Commission, 2013, s.2.
[197] European Commission, 2013, s.12
[198] Rossini ja Green, 2015, s.16.

Henkilötietojen suojan huomioiminen tekoälyjärjestelmissä tulee erityisesti huomioitavaksi tekoälyn opetusvaiheessa datajoukkojen kohdalla sekä tekoälyn käytössä, kun käyttäjät syöttävät tekoälyjärjestelmälle tietojaan. Henkilötietojen käyttö opetusdatana ei ole kiellettyä, mutta tekoälysäädösehdotus toteaa:

> *”Siinä määrin kuin se on ehdottoman välttämätöntä suuririskisiin tekoälyjärjestelmiin liittyvien vinoutumien seurannan, havaitsemisen ja korjaamisen varmistamiseksi, tällaisten järjestelmien tarjoajat voivat käsitellä asetuksen (EU) 2016/679 9 artiklan 1 kohdassa, direktiivin (EU) 2016/680 10 artiklassa ja asetuksen (EU) 2018/1725 10 artiklan 1 kohdassa tarkoitettuja erityisiä henkilötietoryhmiä”.*[199]

Erityisiin henkilötietoryhmiin sisältyvää dataa voitaisiin käyttää suuririskisten tekoälyjärjestelmien opetuksessa harkitusti edellyttäen perusoikeuksien turvaamisen myös teknisesti esimerkiksi pseudonymisoimalla. Lisäksi ehdotuksen artiklassa 54 todetaan myös mahdollisuudesta hyödyntää henkilötietoja toimivaltaisen viranomaisen tai Euroopan tietosuojavaltuutetun perustamien tekoälyn sääntelyn testiympäristöissä tietyin edellytyksin muun muassa, kun kyse on ympäristönsuojelusta tai rikosten ehkäisystä tai tutkinnasta. [200]

5.1 Perusoikeuskirja ja perustuslaki

Euroopan unionin perusoikeuskirjan (2012/C 326/02) kahdeksas artikla määrittelee henkilötietojen suojan olemassaolon. Tuon artiklan mukaisesti:

> *’’Jokaisella on oikeus henkilötietojensa suojaan.’’*

[199] Euroopan komissio, 2021a, artikla 10.
[200] Euroopan komissio, 2021a, artikla 10 ja 54.

Lisäksi artiklan toinen kohta asettaa henkilötietojen käsittelylle tarkoituksenmukaisuuden vaatimuksen sekä käsittelyn edellytykseksi henkilötietojen kohteen suostumuksen tai muun laissa oikeuttavan perusteen.

Edellä mainittujen, hyvin yleisluonteisten, henkilötietojen suojan spesifikaatioiden perintö ilmenee Suomen omassa kansallisessa säädännössä esimerkiksi tietoaineistojen saatavuuden rajoituksissa (Laki julkisen hallinnon tiedonhallinnasta 906/2019 15§); yksityisten ja yhteisöjen tietosuojan huomioimisessa arkistotoiminnassa (arkistolaki 7§) sekä perustuslaissa (731/1999).

Euroopan unionin perusoikeuskirjan sisältö koostuu unionin alueella päteviksi määritellyistä perusoikeuksista. Perusoikeudet ovat siis voimassa jokaisessa EU:n jäsenmaassa laajassa merkityksessään. Vaikka perusoikeuskirja on julkaistu aiemmin jo 2000 luvun alulla, se sai oikeudellisen sitovuuden vasta Lissabonin sopimuksen myötä. Lissabonin sopimuksen kautta näet luotiin sopimus Euroopan unionista (SEU) ja sopimus Euroopan unionin toiminnasta (SEUT). EU- ja EUT-sopimus aloittivat siis Euroopan unionin toiminnan ja kuuluvat EU:n perussopimuksiin aivan kuten Euroopan unionin perusoikeuskirja[201]. Se, miten perusoikeuskirja linkittyy kahteen edellä mainittuun perussopimukseen, perustuu niissä annettuun julistukseen:

> *''Euroopan unionin perusoikeuskirjassa, joka on oikeudellisesti sitova, vahvistetaan ihmisoikeuksien ja perusvapauksien suojaamiseksi tehdyssä eurooppalaisessa yleissopimuksessa taatut jäsenvaltioiden yhteisestä valtiosääntöperinteestä johtuvat perusoikeudet."*

Näin ollen jo pelkästään perusoikeuskirjasta annettuun julistukseen nojautuen, ainakin henkilötietojen suojan voidaan pätevästi määritellä yhdeksi eurooppalaisista perusoikeuksista. Henkilötietojen suoja on tunnustettu myös kansallisella tasolla yhdeksi perustuslaissa ilmenevistä perusoikeuksista osana yksityiselämän suojaa:

---

[201] Ulkoministeriö: Euroopan unionin oikeudelliset kysymykset ja ulkosuhdesopimukset; Eur-Lex: Voimassa olevat perussopimukset.

''Jokaisen yksityiselämä, kunnia ja kotirauha on turvattu. Henkilötietojen suojasta säädetään tarkemmin lailla.'' (Perustuslaki 1999/731 10§).

5.2 Henkilötietojen suoja ja tietosuoja

Erään määritelmän mukaan henkilötieto on tieto, jolla suoraan tai välillisesti pystytään identifioimaan yksilö; tieto, joka sisältää verkkotunnisteen kuten IP-osoitteen, evästeen tai digitaalisen sormenjäljen; sekä paikkatieto, jolla pystytään tunnistamaan yksilö[202]. Käytännössä kaikki tiedot, jotka liittyvät tunnistettuun tai tunnistettavissa olevaan henkilöön ovat henkilötietoja. Samoin tieto, jota yhdistämällä johonkin toiseen tietoon, mahdollistaa tunnistamisen. Henkilötiedoksi luokitellaan muun muassa henkilön nimi, henkilötunnus, jokin tunnusomainen tekijä, potilastieto ja puhelinnumero.[203]

Henkilötietojen ja henkilötietojen suojan lisäksi puhutaan tietosuojasta. Toisin kuin henkilötieto, tietosuoja ei ole konkreettinen elementti. Tietosuoja ennemminkin kuvantaa abstraktia ajatusta tai ohjenuoraa siitä, miten tietoa tai dataa tulisi käsitellä. Sen tarkoituksena on siis osoittaa milloin ja millä edellytyksillä henkilötietoja voidaan käsitellä[204]. Henkilötietojen suoja voidaan taasen nähdä tästä alalohkona käsittäen tietosuojan osalta vain henkilötietoja. Henkilötietojen suojaa ja tietosuojaa saatetaan usein käyttää synonyymeinä keskenään hankkien tälle tavalle hyväksynnän muun muassa yleisestä tietosuoja-asetuksesta, joka datan ja tiedon osalta käsittelee vain henkilötietoja.

Ottaen kuitenkin huomioon erityisesti tietoteknisen kehityksen nopeuden, on selvää, että yhä enenevissä määrin luodaan ja säilötään informaatiota tekniseen muotoon tietokantoihin eikä tuo säilötty data välttämättä ole henkilötietona suojattavaa tietoa, mutta se voi tästä huolimatta tarvita suojattavan tiedon aseman. Edellä mainittu pätee muussakin kuin tietoteknisessä muodossa olevaan ei-henkilötietoon. Henkilötiedon

[202] Goddard, 2017, s.703.
[203] Tietosuojavaltuutetun toimisto: Mikä on henkilötieto?
[204] Tietosuojavaltuutetun toimisto: Tietosuoja.

kategoriasta poikkeavan tiedon tietoturvallisuuden asettamisesta mainitaan kuitenkin Euroopan parlamentin ja neuvoston asetuksessa datahallinnosta[205]. Asetus sisältää ehtoja, joiden mukaisesti välityspalvelujen tarjoajan velvoitetaan toteuttamaan toimenpiteitä, joilla varmistetaan muiden kuin henkilötietojen tallentamisen, käsittelyn ja välittämisen korkea turvallisuustaso (Datanhallinta-asetus art. 12). Tietosuojaa ajatellen onkin, datan ja tiedon osalta, olennaista erotella muukin suojattava informaatio kuin vain henkilötiedot. Henkilötiedot ovat kuitenkin erityisessä asemassa Euroopan unionin perusoikeuskirjan tähden ja vaativat osaltaan laajempaa observointia.

5.3 Tietoturva ja kyberturva

Tietoturvalla viitataan tietosuojan toteuttamisen keinoihin eli tietoturvan tarkoitus on suojata tietoaineistoa ja tietojärjestelmiä esimerkiksi teknisiä ja organisatorisia toimenpiteitä, kuten palomuureja ja kulkukortteja, hyödyntämällä. Erilaiset tietoturvan toimenpiteet varmistavat tiedon luottamuksellisuuden, eheyden ja saatavuuden sekä järjestelmien käytettävyyden ja tiedon kohteen eli rekisteröidyn oikeuksien, muun muassa henkilötietojen suojan sekä tiedon tarkastelemisen, toteutumisen.[206]

Euroopan unionin tietoturvasäädännön osalta on annettu muun muassa direktiivi toimenpiteistä yhteisen korkeatasoisen verkko- ja tietojärjestelmien turvallisuuden varmistamiseksi koko unionissa (2016/1148) eli NIS-direktiivi tai verkko- ja tietojärjestelmädirektiivi[207]. Kyseisellä direktiivillä luodaan tietoturvallisuudelle vähimmäistaso kuitenkin minkään estämättä, että jäsenvaltio voi saavuttaa tätäkin korkeamman verkko- ja tietojärjestelmien turvallisuuden (artikla 3). Lisäksi direktiivi velvoittaa Euroopan unionin jäsenvaltioita hyväksymään verkko- ja tietojärjestelmien turvallisuutta koskevan kansallisen strategian (1 ja 7 artikla).

---

[205] (Euroopan parlamentin ja neuvoston asetus (EU) 2022/868 eurooppalaisen datan hallinnoinnista ja asetuksen (EU) 2018/1724 muuttamisesta (datanhallinta-asetus))
[206] Tietosuojavaltuutetun toimisto: Tietosuoja.
[207] Finanssivalvonta, 2018.

Kansallisella tasolla tietoturvaan viitataan muun muassa laissa sähköisestä asioinnista viranomaistoiminnassa (2003/13 1§):

> *''Tämän lain tarkoituksena on lisätä asioinnin sujuvuutta ja joutuisuutta samoin kuin tietoturvallisuutta hallinnossa, tuomioistuimissa ja muissa lainkäyttöelimissä sekä ulosotossa edistämällä sähköisten tiedonsiirtomenetelmien käyttöä.'',*

Valtioneuvoston asetuksessa asiakirjojen turvallisuusluokittelusta valtionhallinnossa (1101/2019 1§):

> *'' -- sekä turvallisuusluokiteltujen asiakirjojen käsittelyä koskevista tietoturvallisuustoimenpiteistä valtionhallinnon viranomaisissa.'',*

sekä laissa julkisen hallinnon tiedonhallinnasta (906/2019 1§):

> *''Tämän lain tarkoituksena on: 1) varmistaa viranomaisten tietoaineistojen yhdenmukainen ja laadukas hallinta sekä tietoturvallinen käsittely julkisuusperiaatteen toteuttamiseksi''.*

Kyberturvallisuus tulee mainituksi aiemmin käsitellyn verkko- ja tietojärjestelmädirektiivin osalta kyseisen direktiivin ollessa ensimmäinen Euroopan unionin laajuinen kyberturvallisuussäädös. Kyberturvallisuuden voidaan tuon perusteella päätellä käsittävän nimenomaan verkko- ja tietojärjestelmäliikennettä eikä yleisellä tasolla tietoa samalla tavalla kuin tietoturvallisuus. Kyberturvallisuuden määritelmää on pyritty kuitenkin avaamaan tarkemmin kyberturvallisuusasetuksessa (2019/881). Asetuksen toisen artiklan ensimmäisen kohdan mukaan kyberturvallisuudella tarkoitetaan toimia, joita tarvitaan verkko- ja tietojärjestelmien, tällaisten järjestelmien käyttäjien ja muiden asianosaisten suojaamiseksi kyberuhilta. Artiklan kahdeksas kohta paljastaa myös kyberuhan tarkoittavan potentiaalista tilannetta, tapahtumaa tai toimintaa, joka voi vahingoittaa tai häiritä edellä jo mainittuja kohteita.

6 Johtopäätökset

Ihmisen kognitiivisia kyvykkyyksiä ei olla vielä pystytty kopioimaan ja luomaan koneelle eikä ole olemassa ihmisälyä vastaavaa tekoälyä. Tekoäly on tällä hetkellä ”heikkoa tekoälyä”, joka pystyy toimimaan paljon ihmistä rajoitetummin. Tekoälyllä on kuitenkin kyvykkyys oppia itsenäisesti ja tehdä päätöksiä myös ollessaan ennestään tuntemattomien muuttujien kanssa tekemisissä. Näin ollen tekoälylle ei siis ole valmiiksi ohjelmoituna tai opetettuna niitä tiettyjä toimeksiantoja, joita sen halutaan suorittavan vaan tekoäly kykenee soveltamaan aiemmin saadun tiedon avulla. Tekoäly on myös autonominen kokonaisuus ja pystyy toimimaan ilman ihmisen avustusta kuitenkin niin, että autonomisuutta voi esiintyä hyvin eri tasoisena järjestelmästä riippuen.

Kahden edellä mainittujen ominaisuuksien lisäksi tekoälystä on havaittavissa kaksi konkreettisempaa ominaisuutta: data ja algoritmit. Ilman näitä ei voi olla tekoälyä. Tekoäly järjestelmänä koostuu ja sen toiminta ja oppiminen perustuvat koodattuihin algoritmeihin. Oppimista ei lisäksi pystytä mahdollistamaan ilman dataa. Tekoäly vaatii siis dataa ja mitä laajempia datajoukkoja, sitä laajemmin tekoäly myöskin voi oppia ja soveltaa tietoansa erityyppisiin tilanteisiin. Koska tekoäly ei pysty järkeilemään ihmisen tavoin, vaatii tekoäly ihmistä enemmän tietoa saavuttaakseen saman ratkaisun. Tekoäly ei myöskään kyseenalaista toimeksiantoaan tai johda tunteisiin perustuvia ratkaisuja tai päätöksiä.

Vuonna 2018 Euroopan unioni alkoi määritellä tekoälyä julkaisussaan ”Tekoäly Euroopassa” ja myöhemmin Tekoälyä koskevassa valkoisessa kirjassa ja tekoälysäädösehdotuksessaan. Edellä annetusta neljästä ominaisuudesta poiketen unioni ei vaikuta pitävän oppivuutta poissulkevana ominaisuutena. Esimerkiksi Bayesin teoreeman avulla voidaan luoda tekoälyä, vaikkei sen avulla voida oppivuutta luoda. Data, algoritmit ja itsenäisyys ovat kuitenkin unionin määritelmän mukaisesti osa tekoälyä.

Tekoäly ei rajoitu tiettyihin tekniikoihin kuten koneoppimiseen, vaan eurooppalainen määritelmä on haluttu jättää mahdollisimman teknologianeutraaliksi uusien käytettävien tekniikoiden varalle. Lisäksi on määritelty, että tekoälyn ei tarvitse olla itsenäinen

kokonaisuus, vaan se voi olla sulautettu osa tuotetta tullakseen luetuksi tekoälyksi. Käytännössä kaikki Euroopan komission antamat ehdotukset eivät koske kaikkea tekoälyä, vaan erityisesti sääntelyn kohteeksi tulee suuren ja kriittisen riskin aiheuttavat tekoälyjärjestelmät. Näin ollen tekoäly saa uudenlaisen määritystaustan sen luoman riskin perusteella käyttöalan ja käyttötarkoitus huomioiden.

Mitä tulee erityisesti suuririskiseen tekoälyyn, tulee tällaisten järjestelmien täyttää tietyt ehdot, jotka on luotu eurooppalaisia arvoja ja periaatteita noudattaen. Erityisiä suosituksia on annettu AI HLEG-asiantuntijaryhmän toimesta, esimerkiksi eettiset ohjeet luotettavalle tekoälylle ja arviointilista ALTAI. Näiden mukaisesti ihmisen tulisi edelleen hallita tekoälyä joko niin, että ihminen pystyy puuttumaan tekoälyn prosesseihin tai päättämään, ettei tekoälyä käytetä tietyssä tilanteessa, eli ihminen korvaa tekoälyn antaman ratkaisun tai päätöksen. Eurooppalaisen lähestymistavan lähtökohta on, että tekoäly ei ohita ihmisen toimijuutta tai riko perus- ja ihmisoikeuksia ja kunnioittaa yksilöä itsenäisenä toimijana. Käyttäjällä on esimerkiksi aina oikeus tietää, kun hän on tekoälyn kanssa tekemisissä, eikä tekoäly saisi vaikuttaa käyttäjän omaan autonomiaan ja tahtoon.

Tekoälyyn liittyviä muita suosituksia ovat erityisesti datan laadukkuus, ajantasaisuus ja paljous, koska heikot datajoukot voivat johtaa vääristymiin tai virheisiin ja edelleen tiettyjen ihmisryhmien diskriminointiin; sekä tekoälyn prosessien ja päätöksenteon läpinäkyvyys, jotta käyttäjä voi saada selityksen päätökselle. Näin ollen tekoälystä tulisi olla laaja dokumentaatio, jonka avulla voidaan selventää, miten tekoäly on päätynyt tiettyyn ratkaisuun. Dokumentoitavia tietoja on esimerkiksi datan laatu ja käytetyt datajoukot sekä päättelyprosessit, säännöt ja mallit ja, miten ne johtavat tiettyyn tulokseen. Mikäli ei pystytä selittämään ymmärrettävällä tavalla tekoälyn tarjoamia ennusteita ja päätöksiä, kyseessä on ”musta laatikko”, jota tulisi välttää erityisesti julkisella sektorilla tekoälyjärjestelmässä.

Tekoälyn on tietenkin myös aina noudatettava lakia ja standardeja. Näin ollen tekoäly on oltava teknisesti luotettava ja tietoturvallinen sen mukaisesti, mitä säädöksin vaaditaan.

Tekoälyn tulisi olla kestävä erilaisia tietoturvahyökkäyksiä vastaan eikä sen toiminta saisi tällaisten riskien toteutuessa johtaa haitallisiin tai vahingoittaviin seurauksiin. Tekoälyjärjestelmän jokaisen elinkaaren kohdalla on toteuduttava henkilötietojen suoja sen mukaisesti, miten on säädetty kansallisesti ja Euroopan unioni tasolla muun muassa yleisessä tietosuoja-asetuksessa. Yhtenä tekoälyyn liittyvänä riskinä on tunnistettu tekoälyn kyvykkyys luoda itsenäisesti lisää tietoja ja yhdistää dataa henkilöprofiiliksi myös anonymisoidusta tiedosta. Huomiota tulee siten kiinnittää tavallista järjestelmää laajemmin.

Näiden suositusten, ohjeiden, säädösten ja standardien soveltamisen vastuu kuuluu kaikkiin tekoälyn kanssa tekemisissä oleville. Tekoälyn kehitykseen osallistuvilla on vastuu muun muassa luoda laadukas algoritminen pohja sekä valita ja koota datajoukot huolella, tälle tasolle kuuluu myös dokumentaatioiden luonti, eettisten arvojen implementointi järjestelmään ja tietoturvan varmistaminen. Näitä samoja velvollisuuksia kantaa myös tekoälyn tarjoaja. Käyttäjälle itselleen asetettu vastuu nojaa suuresti tarjoajan antamien järjestelmäohjeiden ja tietosuojan noudattamisessa.

Lähteet

Alho, T., Hänninen, P., Neittaanmäki, P. & Tammilehto, O. (2018). Palvelurobotiikka. Jyväskylä: Jyväskylän yliopisto.

Andström K. (2003). Perusasioita oikeustieteestä. Helsingin yliopisto. Noudettu 17.6.2023 osoitteesta: https://www.avoin.helsinki.fi/oppimateriaalit/oikeustiede/materiaali/osa1.html.

BBC. ChatGPT: US lawyer admits using AI for case research. Julkaistu 27.5.2023. Noudettu 12.6.2023 osoitteesta: https://www.bbc.com/news/world-us-canada-65735769

De Bruyne, J., & Vanleenhove, C. (2021). Artificial intelligence and the law. Intersentia.

DeepAI: Weight (Artificial Neural Network). Noudettu 16.4.2023 osoitteesta: https://deepai.org/machine-learning-glossary-and-terms/weight-artificial-neural-network

Eur-Lex: Voimassa olevat perussopimukset. Noudettu 18.9.2021 osoitteesta: https://eurlex.europa.eu/collection/eu-law/treaties/treaties-force.html?locale=fi.

Euroopan komissio (2013). Cybersecurity Strategy of the European Union; An Open, Safe and Secure Cyberspace.

Euroopan komissio (2018). Komission tiedonanto Euroopan parlamentille, neuvostolle, Euroopan talous- ja sosiaalikomitealle ja alueiden komitealle: Tekoäly Euroopassa. COM(2018) 237 final

Euroopan komissio (2020). White Paper on Artificial Intelligence - A European approach to excellence and trust. COM(2020) 65 final

Euroopan komissio (2021a). Ehdotus Euroopan parlamentin ja neuvoston asetus Tekoälyä koskevista yhdenmukaistetuista säännöistä (tekoälysäädös) ja tiettyjen unionin säädösten muuttamisesta. COM(2021) 206 final.

Euroopan komissio (2021b). Komission tiedonanto Euroopan parlamentille, Eurooppaneuvostolle, neuvostolle, Euroopan talous- ja sosiaalikomitealle ja alueiden komitealle: Tekoälyä koskevan eurooppalaisen lähestymistavan edistäminen. COM(2021) 205 final.

Euroopan komissio (2022). Ehdotus Euroopan parlamentin ja neuvoston direktiivi sopimuksenulkoista siviilioikeudellista vastuuta koskevien sääntöjen mukauttamisesta tekoälyyn (direktiivi tekoälyyn liittyvästä vastuusta) COM(2022) 496 final.

Euroopan komissio. A European approach to artificial intelligence. Päivitetty 19.6.2023. Noudettu 16.6.2023 osoitteesta: https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

Euroopan komissio. European data strategy. Noudettu 16.4.2023 osoitteesta: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

Finanssivalvonta (2018). Valvottavatiedote 8.5.2018-30/2018: EU:n Verkko- ja tietoturvadirektiivi kansallisesti voimaan 9.5.2018. Noudettu 18.9.2021 osoitteesta: https://www.finanssivalvonta.fi/tiedotteet-ja-julkaisut/valvottavatiedotteet/20182/eun-verkko--ja-tietoturvadirektiivi-kansallisesti-voimaan-9.5.2018/.

Goddard, M. (2017). The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. International Journal of Market Research, 59(6), 703–705.

Helsingin yliopisto. Elements of AI: Building AI. Verkkokurssi. Noudettu 20.6.2023 osoitteesta: https://buildingai.elementsofai.com/Machine-Learning/linear-regression

Kaarlejärvi, S. & Salminen, T. (2018). Älykäs taloushallinto: Automaation aika. Alma Talent.

Kallioniemi, I. (2022). Tekoälyoikeus. Alma Talent.

Kilpailu- ja kuluttajavirasto. Vahingonkorvaus palvelun virheestä ja viivästyksestä. Noudettu 16.5.2023 osoitteesta: https://www.kkv.fi/kuluttaja-asiat/tavaroiden-ja-palveluiden-virheet/vahingonkorvaus-palvelun-virheesta-ja-viivastyksesta/

Kilpailu- ja kuluttajavirasto. Vahingonkorvaus tavaran virheestä. Noudettu 16.5.2023 osoitteesta: https://www.kkv.fi/kuluttaja-asiat/tavaroiden-ja-palveluiden-virheet/vahingonkorvaus-tavaran-virheesta/

Lappi, O., Rusanen, A.-M., & Pekkanen, J. (2018). Tekoäly ja ihmiskognitio. Tieteessä Tapahtuu, 36(1). Noudettu osoitteesta https://journal.fi/tt/article/view/69278

Massimiliano, Zanin & Nadim, Atiya & Basílio, José & Jan, Baumbach & Benis, Arriel & Behera, Chandan & Bucholc, Magda & Castiglione, Filippo & Chouvarda, Ioanna & Comte, Blandine & Dao, Tien-Tuan & Xuemei, Ding & Pujos-Guillot, Estelle & Filipovic, Nenad & David, Finn & H., Glass & Nissim, Harel & Iesmantas, Tomas & Ivanoska, Ilinka & Schmidt, Harald. (2020). An Early Stage Researcher's Primer on Systems Medicine Terminology.

McDermott, Y. (2017). Conceptualising the right to data protection in an era of Big Data. Big Data & Society, 4(1).

Määttä T., Tolvanen M., Väätänen U., Kolehmainen A., Myrsky M & Keinänen A. (2012). Oikeudellisen ajattelun perusteita: Oikeustieteiden pääsykoekirja 2012. Itä-Suomen yliopisto.

Palkeet. Tekoäly ennustaa ostolaskujen tiliöintejä. Julkaistu 15.9.2021. Noudettu 14.6.2023 osoitteesta: https://www.palkeet.fi/ajankohtaista/tekoaly-ennustaa-ostolaskujen-tiliointeja.html

Rossini, C. & Green, N. (2015). Training the Next Generation of Digital Rights Advocates: Cybersecurity and Human Rights.

Rousku, K. k., Linturi, R., Andersson, C., Stenfors, S. k., Lähteenmäki, I., Kärki, T. & Limnéll, J. (2017). Pilkahduksia tulevaisuuteen: Digitalisaation ja robotisaation mahdollisuudet. Helsinki: Valtiovarainministeriö.

Siukonen, T. & Neittaanmäki, P. (2019). Mitä tulisi tietää tekoälystä. Docendo.

Tekoälyä käsittelevä korkean tason asiantuntijaryhmä (2019a). Policy and investment recommendations for trustworthy AI.

Tekoälyä käsittelevä korkean tason asiantuntijaryhmä (2019b). Luotettavaa tekoälyä koskevat eettiset ohjeet.

Tekoälyä käsittelevä korkean tason asiantuntijaryhmä (2020). The Assesment List For Trustworthy Artificial Intelligence (ALTAI).

Tietosuojavaltuutetun toimisto: Mikä on henkilötieto? Noudettu 18.9.2021 osoitteesta: https://tietosuoja.fi/mika-on-henkilotieto

Tietosuojavaltuutetun toimisto: Tietosuoja. Noudettu 18.9.2021 osoitteesta: https://tietosuoja.fi/tietosuoja

Topol, E. J. (2019). Deep medicine: how artificial intelligence can make healthcare human again. First edition. New York, Basic Books.

Työ- ja elinkeinoministeriö (2017). Työ- ja elinkeinoministeriön julkaisuja: Suomen tekoälyaika: Suomi tekoälyn soveltamisen kärkimaaksi: Tavoite ja toimenpidesuositukset, 41/2017.

Ulkoministeriö. Euroopan unionin oikeudelliset kysymykset ja ulkosuhdesopimukset. Noudettu 18.9.2021 osoitteesta: https://um.fi/eu-oikeus-ja-eu-n-ulkosuhdesopimukset

Valtiovarainministeriö (2017). Sähköisen asioinnin tietoturvallisuus-ohje.

Virtanen, P. & Stenvall, J. (2014). Älykäs julkinen organisaatio. Tietosanoma Oy.

Wikipedia, Vapaa tietosanakirja: Deep learning. Päivitetty 14.4.2023. Noudettu 16.4.2023 osoitteesta: https://en.wikipedia.org/wiki/Deep_learning#Deep_neural_networks

Wikipedia, Vapaa tietosanakirja: Monikerroksinen perseptroniverkko. Päivitetty 3.1.2021. Noudettu 16.4.2023 osoitteesta: https://fi.wikipedia.org/wiki/Monikerroksinen_perseptroniverkko