

Tämä on rinnakkaistallennettu versio alkuperäisestä julkaisusta.

Tämä on julkaisun final draft -versio. HUOM.! Versio voi poiketa alkuperäisestä julkaisusta sivunumeroinnin, typografian ja kuvituksen osalta.

Käytä viittauksessa alkuperäistä lähdettä:

Jalonen, H. 2018. Tekoäly on tekotyhmä ilman eettistä ulottuvuutta. Tietoasiantuntija. Vsk. 33, nro 2 - 3, 20 - 23.

Kaikki julkaisut Turun AMK:n rinnakkaistallennettujen julkaisujen kokoelmassa Theseuksessa ovat tekijänoikeussäännösten alaisia. Kokoelman tai sen osien käyttö on sallittu sähköisessä muodossa tai tulosteena vain henkilökohtaiseen, ei-kaupalliseen tutkimus- ja opetuskäyttöön. Muuhun käyttöön on hankittava tekijänoikeuden haltijan lupa.

This is a self-archived version of the original publication.

The self-archived version is a final draft. NB. The self-archived version may differ from the original in pagination, typographical details and illustrations.

To cite this, use the original publication:

Jalonen, H. 2018. Tekoäly on tekotyhmä ilman eettistä ulottuvuutta. Tietoasiantuntija. Vsk. 33, nro 2 - 3, 20 - 23.

All material supplied via TUAS self-archived publications collection in Theseus repository is protected by copyright laws. Use of all or part of any of the repository collections is permitted only for personal non-commercial, research or educational purposes in digital and print form. You must obtain permission for any other use.

TIETOASiantuntija

33(1-2), 20—23

Harri Jalonen

Tekoäly on tekotyhmä ilman eettistä ulottuvuutta

George Orwell muistetaan parhaiten teoksestaan *Vuonna 1984*. Toisen maailmansodan jälkeen ilmestyneessä ja tulevaisuuteen sijoittuvassa fiktiivisessä romaanissa kuvataan, mitä tapahtuu, kun vallanpitäjillä on käytössään teknologia, joka mahdollistaa alamaisten arjen yksityiskohtaisen seuraamisen ja yleisen mielipiteen tehokkaan muokkaamisen. Orwellin kirjassa kansalaisia seurattiin ja hallittiin heidän olohuoneisiinsa sijoitettujen kaksisuuntaisten teleruutujen välityksellä. Teleruudut olivat pirullisia, sillä ne sekä puhuivat että kuuntelivat ja katselivat. Kehitys johti lopulta tilanteeseen, jossa kansalaisilta oli evätty myös väärin ajattelun mahdollisuus, sillä kielestä oli karsittu vallanpitäjille vahingollisten ajastusten ilmaisuun tarvittavat sanat.

Hyvä, paha tekoäly

Orwellin ja tekoälyn yhdistäminen ei ole erityisen omaperäistä. Esimerkiksi Google antaa huhtikuussa 2018 reilusti toista miljoonaa osumaa sanaparille ”Orwell”–”artificial intelligence”. Pintapuolisen sisällönanalyysin perusteella näyttää siltä, että monissa kirjoituksissa noudatetaan kaavaa, jossa tekoälyn todetaan olevan hyvissä käsissä ihmiskuntaa hyödyttävä teknologia, mutta väärin käsiin joutuessaan seurauksena on orwellilainen dystopia. Tekoälyä on helppo puolustaa, kun se auttaa lääkäreitä diagnosoimaan syöpiä ja muita vakavia sairauksia. Samoin Amazonin asiakkaat kiittävät varastorobotteja, jotka mahdollistavat tuotteiden kustannustehokkaan jakelulogistiikan. Ääni kellossa muuttuu, kun tekoälyn avulla päätellään työntekijöiden sitoutumista firmaan esimerkiksi heidän tietokoneen käyttöaktiivisuutensa perusteella. *The Economistin* (31.3.18) mukaan näin tehdään muun muassa yhdysvaltalaisessa data-analytiikkapalveluja tarjoavassa Veriatossa, jossa rekisteröidään työntekijöiden käyttäytyminen näppäinkomentojen tarkkuudella. Kiinassa puolestaan pilotoidaan pisteytysjärjestelmää (Social Credit Score), jossa kansalaisia lokeroitaan heidän luotettavuutensa perusteella. Parin vuoden kuluttua järjestelmän odotetaan keräävän ja analysoivan informaatiota sadoista erilaisista datalähteistä. Maltillisestikin arvioiden näyttää ilmeiseltä, että kansalaisten saamat luotettavuusluvut vaikuttavat heidän mahdollisuuksiinsa työelämässä ja politiikassa. Harvinaisia eivät ole nekään puheenvuorot, joissa ennakoitaan teknologista singulariteettia eli tekoälyn kehittymistä yli-inhimilliseksi toimijaksi, joka tekee myös vallanpitäjistä palvelijoitaan.

Se, käytetäänkö tekoälyä hyvään vai pahaan, on edelleen ihmiskunnan omissa käsissä. Dystopia ei vaani kulman takana, eikä singulariteettikaan ole tulevaisuutta työkseen ennakoivien ensimmäisillä kalvoilla. Akuutimmat tekoälyn mukanaan tuomat uhkat liittyvät ihmisten yksityisyydensuojaan, algoritmien tekemisiin virhetulkintoihin sekä niiden tuottamaan päätöksenteon vinoumiin. Tekoälyllä on sormensa pelissä myös sosiaalisen median ”kuplissa”, sillä algoritmit filteröivät informaatiota tehokkaasti. Kuplat eivät ole tekoälyn syytä, mutta tekoäly on kärjistänyt some-keskusteluja ja mahdollisesti vaikuttanut myös ihmisten äänestyskäyttäytymiseen.

Tekoäly edellyttää eettistä keskustelua

Tekoäly on teknologiaa, mutta sen haitallisten seurausten minimointi on mitä suurimmassa määrin eettinen kysymys. Näin siksi, koska etiikka käsittelee ensisijaisesti sitä, mikä on oikein, ei sitä, mikä on mahdollista. Tekoälyn juuret ovat 40- ja 50-lukujen taitteessa, mutta etiikkaa pohdittiin jo antiikin Kreikassa. Käsitys oikeasta ja väärästä ei synny yritysten tuotekehitysosastoilla vaan yhteiskunnallisessa keskustelussa. Esimerkiksi konenäön soveltaminen vaikkapa ihmisvirtojen monitorointiin julkisissa tiloissa edellyttää huolellista pohdintaa, jossa katsotaan nykyteknologiaa kauemmaksi sekä arvioidaan erilaisia yhteiskunnallisia kehityskulkuja ja niiden suhdetta teknologiaan. Keskustelu Cambridge Analytica -yhtiön ympärillä osoittaa, että tekoäly mahdollistaa jo nyt asioiden tekemisen tavalla, joka ei kestä päivänvaloa. Riittävä data, tehokkaat algoritmit ja sopiva yhteiskunnallinen tilanne nostavat poliittisen vaikuttamisen riskit uudelle tasolle. Luultavaa on, että tämän päivän täsmäratkaisut näyttävät tulevaisuudessa melko vaarattomalta haulikolla räiskimiseltä.

Etiikassa ollaan kiinnostuneita ihmisten käyttäytymisen moraalisisista periaatteista. Perusoletuksena etiikassa on, että ihminen on tietoinen omasta toiminnastaan ja sen seurauksista. Moraali toimii eräänlaisena kompassina, joka ohjaa eettisesti kestävään toimintaan myös tilanteissa, joihin lainkoura ei yllä.

Tekoälyllä ei ole (ainakaan vielä) tietoisuutta, eikä tekoäly tee robotista ihmistä. Tekoäly voi tehdä eettisesti kestävämpiä ratkaisuja, mutta sitä ei voida itsessään pitää vastuullisena. Siksi eettiset haasteet kulminoituvat siihen, miten sen tekemät päätökset vaikuttavat ihmisiin yksilöinä ja ennen kaikkea ihmisten muodostamiin yhteisöihin.

Tekoälyllä on ulkoisvaikutuksia

MIT:n professori Iyad Rahwanin mukaan tekoälyn eettisten haasteiden jäsentämisessä on syytä korostaa teknologian ja yhteiskunnan suhdetta. Näin siksi, koska tekoäly on teknologiaa, jolla on ulkoisvaikutuksia, jotka kohdistuvat kolmansiin osapuoliin, joilla ei ole ollut sananvaltaa teknologiaa koskevissa päätöksissä. Tekoälyn soveltaminen esimerkiksi itseohjautuvissa autoissa ei voi olla autoteollisuuden sisäinen asia, sillä kysymys on lukuisista liikenteen vaikutuspiirissä olevista sidosryhmistä, joiden odotukset ja vaatimukset tekoälyä kohtaan ovat usein ristiriidassa. Siinä missä yhden painottavat matkanteon vaivattomuutta, ovat toiset huolissaan turvallisuudestaan.

Ihmiset osaavat käyttäytyä huonosti ja siksi meidän toimintaamme on säänneltävä. Sama koskee yrityksiä, joiden toimintaa ohjataan lukuisilla kansallisilla ja yhä useammin kansainvälisillä säädöksillä. Esimerkiksi toukokuussa 2018 voimaantuvassa EU:n tietosuojasetuksessa säädellään periaatteet, joiden mukaan yritykset ja muut organisaatiot voivat käsitellä henkilötietoja. Sääntelyn historia ulottuu kauas ja keinot ovat moninaiset. Hieman yksinkertaistaen kysymys on joko lainsäätäjän määräämistä *laista*, ihmisten odotuksista muodostuvista *normeista* tai kuluttajien ja yritysten toimintaa ohjaavista *markkinoista*.

Tekoäly edellyttää yhteiskuntasopimusta

Eivätkö lait, normit ja markkinat riitä myös tekoälyn sääntelyyn? Periaatteessa kyllä, mutta se edellyttää Rahwanin mielestä tekoälyn erityisluonteen syvällistä ymmärtämistä. Tekoäly ei ole eettis-moraalisiin pohdintoihin kykenevä toimija. Tekoäly ei pode tunnontuskia, eikä iloitse onnistumisista.

Tekoäly ei myöskään ole sorvin kaltainen passiivinen työkalu, joka tarvitsee ihmistä toimiakseen. Tekoälyn tietoisuus on kaukana ihmisestä, mutta tekoälyn kyky oppia kokemastaan tekee siitä aktiivisen toimijan. Tekoäly noudattaa algoritmeja, mutta sen ja ympäristön vuorovaikutus voi johtaa ratkaisuihin, joita kaikkia on vaikea ja jopa mahdotonta ennakoita.

Yhteiskunnallista keskustelua tarvitaan tekoälyyn liittyvien kiperien eettisten kysymysten ratkaisemiseksi, mutta myös tekoälyyn liittyvien perusteettomien pelkojen hälventämiseksi. Monet tekoälyyn liittyvät ongelmat ovat yhteiskunnallisia dilemmoja, joihin ei ole vain yhtä mahdollista vastausta. Kouluesimerkkinä itseohjautuvien autojen käyttäytyminen väistämättömissä onnettomuustilanteissa: pitäisikö auton minimoida kuljettajan vai jalankulkijan vahingot?

Itseohjautuvien autojen tulon valmistaudutaan eri puolilla. Esimerkiksi Saksassa toimii eettinen komissio, joka koostuu eri tieteenalojen asiantuntijoista, teknologian kehittäjistä sekä kansalaisten ja kuluttajien edustajista. Komissiossa kuullaan myös eri uskontokuntien näkemyksiä. Komission tehtävä on kesken, mutta se on jo sopinut muun muassa siitä, että onnettomuustilanteissa tekoälyn ratkaisuun ei saa vaikuttaa osapuolten henkilökohtaiset ominaisuudet, kuten ikä tai sukupuoli. Komissio on myös ohjeistanut, että vahinkojen minimointi osapuolten lukumäärän perusteella on eettisesti hyväksyttävää. Samoin osapuolten oma käyttäytyminen saa vaikuttaa tekoälyn tekemiin ratkaisuihin. Esimerkiksi suojatien ulkopuolella tiellä kävelevän jalankulkijan suoja on huonompi kuin liikennesääntöjä noudattavalla.

Jean-Jacques Rousseau kirjoitti 1700-luvun loppupuolella yhteiskuntasopimuksesta. Sen ydin rakentui eettisesti kestävästä perusteluista kansalaisten ja valtion oikeuksille ja velvollisuuksille. Jotakin samanlaista tarvittaisiin Rahwanin mukaan myös tekoälyn aikakaudella. Yhteiskuntasopimus voi kuulostaa mahtipohtiselta, mutta ei ole sitä enää siinä vaiheessa, kun tekoäly ratkaisee, kuka saa pitää henkensä liikenteessä.

Eettisten valintojen tekeminen ei ole aina helppoa. Kokeile vaikka itse osoitteessa moralmachine.mit.edu

Kirjoituksessa on hyödynnetty MIT Sloanin ja MIT CSAILin kurssia Artificial Intelligence: Implications for business strategy.

Acknowledgements

This paper was created in conjunction with the Co-creation of Service Innovation in Europe (CoSIE) project. The project has received funding from the European Union's Horizon 2020 research and innovation programme H2020-SC6-COCREATION-2017 under grant agreement No 770492. www.cosie-project.eu

Publication date: June 4, 2018