YHTEISKUNNALLISESTI KESTÄVÄ TEKOÄLY

Petri Ylikoski

Professori, VTT

Helsingin yliopisto

Sisällys

Mistä on kyse? 3

Kestävän tekoälyn edellytys 4

Luotettava data 4

Datan laatu 4

Datan turvallisuus 4

Vinoutunut data 5

Datatyö 6

Data ja kannustimet 5

Datalukutaito 6

Viestit päättäjille ja kansalaisille 7

Politiikkasuositukset 7

Datatyön kustannusten huomioonottaminen 7

Datalukutaidon koulutus ..... **Virhe. Kirjanmerkkiä ei ole määritetty.**

Koko järjestelmän elinkaaren kattava läpinäkyvyys ja vaikutusten arviointi ..... **Virhe. Kirjanmerkkiä ei ole määritetty.**

Seuraavilla sivuilla on muutamia saavutettavuuteen liittyviä erityisohjeita ..... **Virhe. Kirjanmerkkiä ei ole määritetty.**

Hyperlinkkien merkitseminen ..... **Virhe. Kirjanmerkkiä ei ole määritetty.**

Alaviitteiden käyttö ..... **Virhe. Kirjanmerkkiä ei ole määritetty.**

Työ voi alkaa ..... **Virhe. Kirjanmerkkiä ei ole määritetty.**

Mistä on kyse?

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

Viimeaikaiset tekoälyn kehitysaskeleet ovat perustuneet erityisesti laskentatehon ja erilaisten koneoppimistekniikoiden kehittymiseen. Vaikka koneoppimismallit tuottavatkin joitakin erityisiä haasteita, on hyvä pitää mielessä, että monet tekoälyjärjestelmien tuottamat haasteet ovat tuttuja jo vanhemmista tieto- ja tietokonejärjestelmistä. Tällaiset järjestelmät ovat usein vuosia kestäneen jatkuvan kehitystyön seurauksena hyvin monimutkaisia ja siksi niiden luotettavuuden arviointi ja toiminnan ymmärtäminen on usein vaikeaa, mikä tuottaa ongelmia myös niiden hallittavuudelle. Tekoälyelementtien lisääminen ei hävitä näitä haasteita, vaan pikemminkin kärjistää niitä.

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

Yleistä keskustelua tekoälystä ohjaa odotus, jonka mukaan tekoälyjärjestelmille kehittyy pian ihmisenkaltainen järkeilykyky, tietoisuus ja moraalinen toimijuus. Tällaisen odotuksen ohjaamana esimerkiksi sanomalehdissä kirjoitetaan tekoälystä kuin se olisi jo nyt toiminnan subjekti, joka kykenee tekemään ennusteita ja vaikkapa tieteellisiä löytöjä. Tämä on tekoäly mystifiointia ja se tekee lukijoille vaikeaksi ymmärtää mistä koneoppimisen soveltamisessa on oikeastaan kysymys. Se luo myös vääränlaisen uhkakuvan. Tekoälyn kehittymisen suurin uhka ei ole se, että ihmistä huomattavasti älykkäämmät ja viisaammat koneet tulevat hallitsemaan elämäämme. Tämä on nykytiedon valossa fantasiaa. Huomattavasti välittömämpi uhkakuva on sellainen, jossa yhteiskuntamme on täynnä epämääräistä dataa käyttäviä ja huolimattomasti suunniteltuja tekoälyjärjestelmiä, joiden käyttäjät eivät ymmärrä niiden toimintaa. Tällöin eläisimme järjestelmien armoilla ja kenelläkään ei olisi kontrollia. Tällaisen tilanteen välttäminen yhteiskunnallisesti kestävän tekoälyn keskeinen haaste.

Kestävän tekoälyn edellytys

Luotettava data

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

Datan laatu

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

Datan turvallisuus

Digitalisoituva maailmaa mahdollistaa yhä laajemman yksilöitä koskevan datan keräämisen ja yhdistelemisen. Tätä dataa voidaan käyttää monella tapaa yksilöiden ja yhteiskuntien hyödyksi. Kerättyyn dataan sisältyy kuitenkin myös merkittäviä vaaroja, sillä väärin käsiin joutuessaan – tai yksinkertaisesti väärin käytettynä – se uhkaa sekä yksilön yksityisyydensuojaa että turvallisuutta. Tästä syystä kansalaisten luottamus datan käyttöä koskevaan säätelyyn on tärkeää. On syytä tähdentää, että kysymys ei ole vain uskomuksesta, että data on turvassa. Jotta datan käyttö olisi kestävällä pohjalla, tulisi järjestelyjen toimia siten ne antavat kansalaisille hyvät perusteet luottamukselleen. Heikko luottamus datan turvallisuuteen tai datan käyttäjien tarkoitusperiin on merkittävä este yhteiskunnallisesti kestävälle tekoälyn soveltamiselle. Lisäksi se saattaa johtaa käytettävissä olevan datan laadun heikkenemiseen, kun kansalaiset kieltävät

datansa käytön, eivät luovuta sitä lainkaan tai vääristävät dataansa. Käytännön esimerkki tästä on Internet-käyttäjien asenne monien sivustojen keräämiin henkilötietoihin.

Vinoutunut data

Automaattisessa päätöksenteossa voidaan erotella sääntöpohjaiset ja koneoppimiseen perustuvat ratkaisut. Sääntöpohjaisissa järjestelmissä keskeisiä riskejä ovat virheellisestä datasta tai ohjelmointivirheistä aiheutuvat virheelliset päätökset. Sen sijaan koneoppimiseen perustuvissa järjestelmissä on kaksi lisäongelmaa. Ensimmäinen näistä on se, että mikäli järjestelmän opettamisessa käytetty data on vinoutunutta, esimerkiksi syrjivien käytäntöjen tuloksena, saattaa järjestelmä oppia nämä samat vinoumat ja ryhtyä toistamaan vastaavaa syrjivää käyttäytymistä. Pelkkä datan laatu ja edustavuus ei auta tällaisessa tilanteessa. Tilanteen diagnosointi on vaikeaa, sillä tässä vaiheessa astuu kuvaan toinen tärkeä koneoppimismallien ominaisuus, niiden läpinäkymättömyys. Tämä tarkoittaa sitä, että täsmälliset periaatteet, jotka malli on itselleen omaksunut ovat käyttäjälle vaikeita tai mahdottomia ymmärtää. Mallia voidaan arvioida lähinnä sen seurauksien kannalta: kuinka se käsittelee erityyppisiä esimerkkitapauksia. Tällainen testaaminen vaatii aikaa, mutta myös käsityksen siitä millaisia arvostelmia järjestelmän pitäisi antaa. Toisin sanoen, järjestelmän testaamisessa ja sen mahdollisten vinoumien korjaamisessa tarvitaan käsitys siitä millainen maailma olisi ilman ei-toivuttua syrjintää. Tämä vaatii laajempaa osaamista kuin data-analyytikon perustaidot. Kestävä tekoälyn soveltaminen yhteiskuntaan edellyttää datan rajoitusten tunnistamista.

Data ja kannustimet

Yhteiskuntatieteissä on jo pitkään tunnettu periaate, jonka mukaan silloin kun mittarista tulee päämäärä, se lakkaa olemasta hyvä mittari. Tämä tarkoittaa sitä, että kun mittarit, joilla arvioidaan organisaation onnistumista päämääriensä saavuttamisessa, asetetaan palkkioperusteiksi organisaatiossa toimiville yksilöille tai ryhmissä, alkaa prosessi, jonka seurauksena organisaation toiminta vääristyy ja alkuperäinen mittari ei enää kuvaan todellista onnistumista päämäärien tavoittelussa. Mittarin käyttö palkitsemisen perusteena siis tuottaa vääränlaisia kannustimia, joiden seuraaminen saattaa johtaa organisaation toiminnan vaarantumiseen. Palkkiojärjestelmien suunnittelijoiden on syytä tuntea kuinka käytetyt mittarit toimivat. Tämä vanha viisaus on syytä muistaa myös tekoälyn aikaudella, sillä jatkuva datafikaatio ja analytiikkaohjelmistojen kehittyminen luo jatkuvasti uusia mahdollisuuksia palkkio- ja seurantajärjestelmille, jotka eivät välttämättä palvele organisaation päämääriä. Tekoäly ei ymmärrä organisaation toimintaa ja ihmiset ovat erittäin taitavia oppimaan kuinka konetta voi huijata. Esimerkiksi järjestelmä, joka arvioi asiakaspalvelun ystävällisyyttä mittaamalla henkilökunnan hymyilyä saattaa nopeasti luoda kummallista käyttäytymistä, jos kasvojentunnistusohjelman tunnistamat ja pisteyttämät hymyt eivät ole erityisen hyvin kytköksissä asiakkaiden kokemuksiin. Tällainen tekoälyn ohjaama mikromanageroitu sosiaalinen vuorovaikutus, josta on jo kokeiluja ulkomailta, tuskin tuottaa työoloja, jotka ovat

yhteiskunnallisesti kestäviä. Tekoälyn sovellutuksia tuleekin aina arvioida osana laajempaa kokonaisuutta, jotta niiden vaikutukset inhimilliselle vuorovaikutukselle ja organisaatioiden toiminnalle tulee selväksi.

Datatyö

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

Datalukutaito

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

Datalukutaitoa voidaan ajatella sekä kaikkien jakamana kansalaistaitona että määrättyjen tehtävien edellyttäminä taitoina. Jälkimmäisessä merkityksessä voimme pohtia millaista datalukutaitoa rekisteriä ylläpitävä viranomainen, yhteiskuntatieteilijä, valmisteleva virkamies tai poliittinen päättäjä tarvitsee. Datalukutaitoa tarvitsevat myös yritysjohtajat, terveydenhuollon ammattilaiset ja kunnalliset päättäjät. Datalukutaito on edellytys realistisille odotuksille koskien uusia data-analytiikka menetelmiä. Se mahdollistaa myös järkevämmän keskustelun tekoälyn avaamista yhteiskunnallisista mahdollisuuksista ja sen luomista vaaroista. Yhteiskunnallisesti kestävä tekoäly ei edellytä vain tekoälymenetelmiin erikoistuneiden asiantuntijoiden ja tutkijoiden kouluttamista, mutta myös huomattavasti laajempaa dataan ja käyttöihin liittyvän ymmärryksen kehittämistä niin kansalaisten kuin päätöksentekijöidenkin keskuudessa.

Viestit päättäjille ja kansalaisille

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

Politiikkasuositukset

Datan laadun merkitys

Suunniteltaessa uusia tekoäly- tai tietojärjestelmiä on otettava huomioon tarkoituksenmukaisen ja laadukkaan datan saatavuus, sekä sellaisen luomisen kustannukset. Tämä antaa paremman kuvan hankkeen kokonaiskustannuksista ja antaa paremmat mahdollisuudet hankkeen onnistumiselle.

Datatyön huomioonottaminen

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

Datalukutaidon koulutus

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