

Asia: VM057:00/2018 VM/2527/00.01.00.01/2017

Luonnos valtioneuvoston selonteoksi "Eettistä tietopolitiikkaa tekoälyn aikakaudella"

Lausunnonantajan lausunto

Kirjoittakaa luonnosta koskevat yleiset kommenttinne tähän:

CSC - Tieteen tietotekniikan keskus Oy (CSC) kiittää mahdollisuudesta saada lausua luonnoksesta valtioneuvoston selonteoksi Eettistä tietopolitiikkaa tekoälyn aikakaudella. CSC on suomalainen ICT-osaamiskeskus, joka ylläpitää opetus- ja kulttuuriministeriön toimeksiannosta korkeakoulujen valtakunnallista keskitettyä tietotekniikkainfrastruktuuria, ja tarjoaa sen avulla kansallisia tietotekniikkapalveluita tutkimuksen, koulutuksen, kulttuurin ja julkishallinnon tarpeisiin.

Selonteon aihe on hyvin ajankohtainen ja tärkeä. Tekoälyn etiikkaan liittyvien kysymysten on arvioitu olevan jopa ihmiskunnan tämän hetken tärkeimpien haasteiden joukossa, ilmastonmuutoksen ja ydinaseiden rinnalla.

Tekoäly on teknologia, jonka täysimääräinen hyödyntäminen vaatii datan vapaampaa liikkuvuutta ja suurta laskentatehoa. Tarvitaan myös strategista ajattelua sen suhteen, mihin suuntaan tekoälyn keskiössä olevan datan käyttöä ohjataan. Suomeen tulee laatia strateginen, poikkihallinnollinen tekoälyn tilannekuva, jossa määritellään tavoitteet ja arvioidaan osaamis- ja resurssitarpeet. Samoin tulee edistää eettisesti kestävä ja teknisesti korkeatasoinen tekoäly- ja dataekosysteemin rakentamista. On myös huolehdittava kestävästä panostuksista datanhallintaan, laskentaan ja datan eettiseen käyttöön. Suomesta on tehtävä datan eettisen hyödyntämisen edelläkävijä.

Tekoäly tarjoaa alati kasvavia mahdollisuuksia, ja sitä käytetään tulevaisuudessa entistä useammassa yhteyksissä ja sovelluksissa. Tekoälyn avulla voidaan ratkaista monimutkaisia ongelmia laskennallisin menetelmin, ja se voidaan ohjelmoida oppimaan tekemään itsenäisiä ratkaisuja sille annetun datan avulla. Kuitenkin ihminen määrittelee, millaiset valtuudet tekoälyllä on ja millä reunaehdoilla se voi tehdä tai oppia tekemään päätöksiä. Tekoäly sinänsä on arvoneutraalia, ja sen eettisyyden määrittävät käyttötarkoitus ja -tapa. Tekoälyn avulla voidaan esimerkiksi parantaa julkisia palveluja,

mutta sitä voidaan käyttää myös esimerkiksi kansalaisten kaiken kattavaan valvomiseen ja kontrollointiin perusoikeuksia loukaten.

CSC:n mielestä on hyvä, että selonteko on valmisteltu laajapohjaisessa yhteistyössä eri tahojen kanssa ja erilaisia näkemyksiä huomioiden. Eri lukujen näkökulmien moninaisuus kuitenkin aiheuttaa hajanaisuutta, joka vaikuttaa selonteon luettavuuteen. Selonteko hyötyisi toimitustyöstä, jossa sen eri luvut pyrittäisiin liittämään kiinteämmin toisiinsa, jotta niistä muodostuisi yhtenäinen näkemys. Selonteossa halutaan esimerkiksi edistää sekä yritysten menestystä että kansalaisten tietosuojaa eettisen tietopolitiikan avulla. Tällä hetkellä ainakin monopolististen kuluttajapalveluiden (tyyppiesimerkkinä Google) kohdalla tietosuoja ja kaupallinen menestys ovat käytännössä ristiriidassa. Selonteosta ei löydy perusteluja, minkälaiset kannustimet voisivat johtaa tämän ristiriidan purkamiseen. Miksi eettisyyden lisääminen parantaisi yritysten kilpailukykyä?

Yksi mahdollinen selontekoa yhtenäistävä tarina voisi olla ottaa lähtökohdaksi luvussa 3.3 esitetty visio Hyvän elämän Suomesta. Tekniset ja kaupalliset intressit ovat usein dominoineet tekoälyn kehityksessä ja eettisiin ongelmiin on havahduttu vasta, kun ongelmia on syntynyt. Selonteon keskeinen näkökulma on Suomen erottuminen kansainvälisessä tekoälykilpailussa eettisellä ja ihmisläheisellä lähestymistavalla. Siksi olisi luontevaa, että selonteko etenisi eettinen näkökulma edellä määrittelemällä ensin, mitä hyviä päämääriä halutaan edistää ja sen jälkeen pohtimalla, miten ne toteutetaan tekniikan ja tietopolitiikan keinoin.

Suomen tulee olla aktiivisena toimijana valjastamassa tekoälyn yhteiskunnalliset vaikutukset teknologisesti korkeatasoisella ja eettisesti kestäväällä tavalla. Tekoälyn kehittäminen vaatii kestäviä panostuksia tehokkaaseen laskentaan, ja tekoälyn raaka-aineena oleva data on läpileikkaava ilmiö, joka ei tunne hallinnonalojen tai sektoreiden rajoja. Siksi Suomen tulee paitsi turvata kansalliset laskentaresurssit, myös järjestelmällisesti poistaa datan liikkuvuuden esteitä, jotta Suomi ei jää jälkeen kehityksestä. Jatkuva oppiminen ja osaamisen kehittäminen ovat elintärkeitä Suomen kilpailukyvyn kannalta.

Pienenä huomiona todettakoon, että selonteossa esiintyy epä johdonmukaisuutta käsitteiden "tieto" ja "data" käytössä (esimerkiksi Tiivistelmässä). Toimiakseen tekoäly ei välttämättä tarvitse tietoa (prosessoitua dataa), vaan sitä voidaan käyttää myös uuden tiedon tuottamiseksi käsittelemättömästä raakadatasta.

CSC:n mielestä tiettyjä näkökulmia tulee huomioida selonteossa vielä tarkemmin, ja näihin ehdotuksiin palataan myöhemmin tässä lausunnossa.

Kirjoittakaa kohtaa 1 koskevat kommenttinne tähän:

Selonteossa keskeistä tietopolitiikan määritelmää tulee täsmentää tarkentamalla tarvittavat "politiikkatoimenpiteet" esimerkiksi "valtion toimiksi, joiden tarkoituksena on edistää tiedon hyvää hallintaa ja tehokasta hyödyntämistä". Tekoölyn yhteys tietopolitiikkaan jää johdannossa hieman ohueksi. Ilmeisesti tekoöly nähdään suhteellisen itsenäisenä osana tiedon hyödyntämistä, vaikka tulevaisuudessa se tulee olemaan entistä olennaisempi tiedon hyödyntämisen keino

CSC pitää hyvänä, että selonteon kuvaus tekoälystä on yhteensopiva VNK:n kesäkuussa julkaiseman raportin Tekoölyn kokonaiskuva ja osaamiskartoitus (<https://tietokayttoon.fi/julkaisut/raportti?pubid=URN:ISBN:978-952-287-549-5>) kanssa, jonka tutkijajäsenet toivoivat, että kyseinen raportti toimisi valtionhallinnossa pohjatyönä ja yhteisenä näkökulmana.

Kuvassa 1 Yhteiskunta on esitetty rinnakkaisena toimijana yksilöille ja yrityksille. Yleensä käsitettä "yhteiskunta" pidetään yläkäsitteenä, joka koostuu ihmisistä ja heidän yhteisöistään, sisältäen muun muassa yksilöt ja yritykset. Yhteiskunnan sijasta kuvan kolmas pilari voisi olla esimerkiksi julkishallinto.

Teknologia sinänsä ei ole hyvää tai pahaa, mutta välineenä se voi mahdollistaa asioita, joita sen hetkessä tai myöhemmässä tarkastelussa voidaan pitää pahoina tai ei-toivottavina. Selontekoon liittyvässä yhteiskunnallisessa keskustelussa valtion pitää haastaa mukaan myös kansalaisyhteiskunta ja kansainväliset toimijat.

Kirjoittakaa kohtaa 2 koskevat kommenttinne tähän:

Luvussa todetaan Yhdysvaltojen, Kiinan ja Japanin hegemonia datataloudessa. Euroopassa painotetaan eettisten periaatteiden huomioimista tekoölyn hyödyntämisessä. Monet valtiot pyrkivät tekoölyinnovaatioiden suunnannäyttäjäksi, mutta epäselväksi jää, mitä sillä tarkoitetaan. Arvioidaanko edelläkävijyyttä teknisten ominaisuuksien, taloudellisten, vai kenties eettisten mittarien kautta? Miten ihmisläheisyys ja eettinen tekoöly auttavat Eurooppaa ja Suomea parantamaan kilpailukykyään?

Kansainvälistä tilannetta tarkasteltaessa on aina muistettava erityisesti yritysten osalta, ettei pääomalla ole kotimaata tai -maanosaa. Erityisesti digitaalisella aikakaudella toiminta voidaan helposti siirtää sinne, missä toimintamahdollisuudet esimerkiksi regulaation suhteen ovat suotuisimmat. Toisaalta yritykset voivat toimia eettisesti korkeatasoisesti myös sellaisissa olosuhteissa, joissa niiltä ei varsinaisesti sitä vaadita. Tämä johtaa usein kilpailijoita suurempiin kustannuksiin ja siten markkina-aseman heikkenemiseen. Kuitenkin vastuullinen ja eettinen toiminta voi nykyisin olla tiedostavien asiakkaiden ansiosta kilpailuvaltti. Silti epäeettiseksi tai epäreiluksi koettu toiminta ei läheskään aina saa kuluttajaa luopumaan palvelun käyttämisestä edes silloin, kun hän itse mahdollisesti kärsii vahinkoa. Esimerkkejä tästä löytyy etenkin sosiaalisen median suuryrityksistä.

Tekoälyn eettisistä periaatteista sopiminen kansainvälisellä tasolla on välttämätöntä, jotta pelisäännöt ovat kaikilla yhteiset. Suomen tulee olla aktiivinen eri toimijoissa ja pyrkiä edistämään selonteossa ehdotettuja linjauksia ja politiikkoja niiden edistämiseksi. Keskustelua tekoälyn periaatteista käydään valtioiden ja niiden välisten organisaatioiden lisäksi muillakin foorumeilla, joihin on myös mentävä rohkeasti, näistä esimerkiksi Partnership on AI -verkosto.

Kirjoittakaa kohtaa 3 koskevat kommenttinne tähän:

Luvussa on ansiokkaasti analysoitu Suomen tilannetta vahvuuksineen, haasteineen, mahdollisuuksineen ja uhkineen. Suomessa on perinteisesti korkea luottamus hallinnon ja viranomaisten edustajiin. Viime aikoina on tullut ilmi valitettavia väärinkäytöksiä, kuten tietosuojarikkomuksia. Tekoäly mahdollistaa myös sellaista toimintaa, jota ei voida pitää eettisenä. Luottamus on helppo menettää nopeasti, ja sen takaisin saaminen voi olla vaikeaa, ellei mahdotonta. Siksi kaiken tekoälyn hyödyntämisen julkisessa toiminnassa tulee olla läpinäkyvää ja perustua yhteisesti sovittuihin, esimerkiksi perustuslaissa määriteltyihin, periaatteisiin ja arvoihin. Tekoälyä on mahdollista käyttää väärin, ja tällaisen toiminnan varalta on oltava toimiva seuranta- ja sanktiointijärjestelmä.

Tekoälyn mahdollisuuksia ja käyttöä julkisessa hallinnossa voisi selonteossa käsitellä laajemminkin. Lähes puuttuva näkökulma on, miten tekoälyä voi hyödyntää päätöksenteon tukena.

Kappaleessa 3.3 kuvattu visio hyvän elämän Suomesta voisi olla laajempi ja selkeämmin selonteon kantavana kertomuksena. Vision pohjalta tulee avata miten tiedon hyvä hallinta ja tehokas hyödyntäminen tukevat halutun vision toteuttamista. Lisäksi CSC:n mielestä tekoälyn etiikan arvioinnissa pitää kiinnittää huomiota siihen, kuka hyötyy tällaisesta päätöksestä, mihin tietoa käytetään ja kenen intressejä palvelee.

Miten Suomesta tehdään datan ja tekoälyn eettisen hyödyntämisen edelläkävijä? Eettisesti kestävä ja teknisesti korkeatasoinen tekoäly- ja dataekosysteemin rakentaminen vaatii panostuksia laskentaresursseihin sekä osaamisen kehittämistä. Laskentaresursseja tarvitaan yhteiskunnan eri sektoreilla, koska suurten datamäärien hyödyntäminen ja tekoälyn kehittäminen eivät onnistu ilman tieteellistä laskentaa, jolla dataa analysoidaan ja jatkojalostetaan sellaiseen muotoon, joka on yhteiskunnan ja elinkeinoelämän hyödynnettävissä. Suuri osa eri alojen huippututkimuksesta tarvitsee suurteholaskentaa. Laskentakapasiteettia tarvitaan esimerkiksi merkittävien globaalien haasteiden ratkomisessa, kuten ilmastomallinnuksessa ja lääketieteessä.

Tekoälyn kehittämisen tärkeä edellytys on lainsäädännöllisten esteiden purkaminen. Hyvin suunnitellun sääntelyn ja sopimusten myötä henkilökohtaisen ja teollisuuden datan hyödyntäminen yhdessä julkisesti saatavilla olevan datan kanssa synnyttävät parhaimmillaan datamarkkinat, joista kaikki hyötyvät. Eri sektoreiden – kaupallinen, tutkimus, julkinen – välistä datan liikkuvuutta,

yhdistelyä, uudelleenkäytettävyyttä ja löydettävyyttä tulisi Suomessa aktiivisesti kehittää. Tässä on tärkeää soveltaa niin sanottua eurooppalaista yhteentoimivuusviitekehystä (European Interoperability Framework) kaikilla sen tasoilla, lainsäädännöllisten esteiden kartoittamisesta ja purkamisesta aina teknisiin sovelluksiin. Tieteen ja liike-elämän yhteinen kehitystyö vaatii datan liikkuvuutta eri sektoreiden välillä, ja siksi on erityisen tärkeää myös Suomen tasolla purkaa hallinnonalojen väliset siilot, jotta päästään aitoon dialogiin ja yhteistyöhön sekä voidaan hyödyntää datan koko potentiaali.

CSC:n mielestä on erittäin hyvä, että selonteossa on huomioitu ihmisten oikeus omiin tietoihinsa omadata –ajattelun mukaisesti ja kansalaisten perusoikeudet tekoälyn kehittyessä.

Kirjoittakaa kohtaa 4.1 koskevat kommenttinne tähän:

Luvussa on ansiokkaasti kerätty toimintasuosituksia. Oikeakielisyyden näkökulmasta CSC ehdottaa otsikon muotoiluksi ”Meidän tiemme”.

CSC tukee tiedon ja datan mahdollisimman laajaa avaamista ja tehokasta käyttöä. Kappaleessa esitetyn lisäksi tulee huomioida, että data ja data-analytiikka ovat yhä keskeisempi osa tutkimusta, innovaatioita, liiketoimintaa ja julkisen sektorin palvelutuotantoa. Laskennallista data-analyysia hyödynnetään enenevässä määrin muun muassa mallintamisessa ja ennusteiden luomisessa. Tekstin- ja datanlouhintateknologia on yleistä kaikkialla digitaalitaloudessa, ja yleisesti tunnustetaan, että tekstin- ja datanlouhinta on erityisen hyödyllistä tutkimusyhteisöille ja datapohjaisten palveluiden kehittämiseksi sekä muun muassa alustatalouden ja tekoälyn luomiselle, joka vaatii suurien datamassojen analysointia ja tulkintaa.

Suomen tuleekin systemaattisesti tukea sitä, että tekstin- ja datanlouhinta on kansallisella ja kansainvälisellä säädösten tasolla mahdollista sekä tieteellisiin että kaupallisiin tarkoituksiin, jotta voidaan hyödyntää datassa oleva arvo liiketoiminnan luomisessa ja edistämiseksi sekä eurooppalaisen datatalouden rakentamisessa. Jos näin ei tehdä, datatalouden kehittäminen vaarantuu. Suomen tulee, tekijänoikeudet tarpeellisessa määrin huomioiden, ajaa aktiivisesti EU:ssa ja kansallisesti mahdollistavaa sääntelyä, jossa datanlouhintaa ei rajoiteta turhaan.

Kirjoittakaa kohtaa 4.2 koskevat kommenttinne tähän:

Luku sisältää erinomaista pohdintaa ja on koko selonteon keskeisintä antia. Huomio siitä, ettei eettisyys ole luonnontieteellisesti määriteltävissä, on tärkeä, ja sitä koskevaa pohdintaa tulee laajentaa. Samoin lukuun tulee lisätä konkretia kuvaamalla tarkemmin, mitä soveltaminen tarkoittaa ja mihin kohtaan järjestelmää eettisen pohdinnan kärki tulisi kohdistaa.

Luvussa tulee myös hyvin esille eettisten kysymysten perustavanlaatuisuus. Tietopolitiikassakin tulee edetä visio edellä – ensin määritellä tahtotila ja vasta sen jälkeen pohtia, miten olemassa olevaa tekniikka voidaan hyödyntää sen toteuttamiseen. Käytännössä tilanne on usein juuri päinvastainen: teknologia kehittyy sääntelyä nopeammin ja mahdollistaa asioita, jotka hetkellisesti ovat ehkä sallittuja, mutta joita ei voida pitää moraalisesti kestävinä.

Kirjoittakaa kohtaa 4.3 koskevat kommenttinne tähän:

Koneelta puuttuu ihmiselle olennaisesti kuuluva inhimillinen ulottuvuus ja kulttuurinen konteksti, joka määrittelee eettisen ja moraalisesti oikean toiminnan. Ihmistä ohjaavat paitsi tieto ja annetut ohjeet, myös osittain sisäsyntyinen ja osittain opittu käsitys oikeasta ja väärästä. Tekoälyn ”moraali”, jos sillä koskaan sellaista voidaan katsoa olevan, on aina ihmisen määrittelemää.

Datan käytössä voidaan erottaa neljä vaihetta: datan kerääminen, kuvailu, uuden tiedon löytäminen ja lainalaisuuksien löytäminen, tiedon perusteella luotavat ennusteet, ja viimeisenä datan perusteella muodostettavat ja annettavat toimitasuositukset. Tekoäly kuuluu varsinaisesti vasta viimeiseen vaiheeseen, jossa se käsittelemänsä datan perusteella antaa toimitasuosituksia. Etiikka ja moraalit liittyvät olennaisesti tähän vaiheeseen. Tapauksesta riippuen suositukset voivat olla eettisesti ongelmattomia, esimerkiksi optimoitu reittiohje paikasta toiseen. Tilanne on toinen, jos tekoälyn antamat suositukset liittyvät esimerkiksi ihmisen terveyteen tai muihin henkilökohtaisen elämän piiriin kuuluviin asioihin. Eettinen harkinta sekä ihmisen ja koneen vuorovaikutukseen kohdistuvat toimenpiteet pitää kohdentaa juuri tälle suositustasolle.

Kirjoittakaa kohtaa 4.4 koskevat kommenttinne tähän:

Tekoälyn mahdollisimman kattava hyödyntäminen vaatii olemassa olevan suomalaisen osaamisen jatkuvaa kehittämistä. Alan koulutusta on myös merkittävästi lisättävä kaikilla asteilla. Koulutusta tulee kehittää yritysten, tutkimus- ja koulutuslaitosten sekä julkisten organisaatioiden yhteistyöllä. Tekoälyn avulla voidaan kehittää esimerkiksi oppimisanalytiikkaa ja oppimista tukevia sovelluksia, joiden avulla oppija voi edetä juuri hänelle sopivalla tavalla omaa henkilökohtaista opintopolkuaan.

Data-analyysi on keskeinen osaamisalue kasvun ja innovoinnin kannalta ja siksi onkin tärkeää, että datatieteen (data science) koulutus tuodaan useille eri sovellusalueille, jotta yhteiskunnan palvelukseen saadaan monipuolisia ja datankäsittelyn haasteisiin vastaavia osaajia. Erityisen tärkeää on huomioida, että eri aloilla tarvitaan erilaista dataosaamista. Digitaalisten taitojen suhteen tulee myös varmistaa, että kansalaisilla on tarvittava osaaminen datan arvon ymmärtämiseen: miten dataa käytetään ja mitkä ovat yksilön oikeudet omien tietojensa suhteen. Tämä on oleellista myös omadata –ajattelun näkökulmasta. Dataosaamista ja -taitoja tulisi myös sisällyttää elinikäisen oppimisen avaintaitojen määritelmään.

Dataosaamisen ja -taitojen olennaisena osana tulee olla myös ymmärrys datan hyödyntämiseen liittyvistä eettisistä kysymyksistä. Tämä tulee huomioida paitsi koulutusohjelmien suunnittelussa,

myös osana digikansalaisen yleissivistystä. Lisäksi tiedollisten taitojen lisäksi tulevaisuudessa tarvitaan entistä enemmän myös itsetuntemusta, koska se on osa kriittistä medialukutaitoa ajassa, jossa eri intressitahojen vaikuttamisyritykset ovat tehokkaampia kuin koskaan aikaisemmin.

Kirjoittakaa kohtaa 4.5 koskevat kommenttinne tähän:

Yleisesti datatalouden kehittämisen pitää tarkoittaa sitä, että tuetaan aidosti datapohjaisia palveluita kehittäviä startup- ja kasvuyrityksiä ja luodaan niiden liiketoiminnalle suotuisat ja houkuttelevat olosuhteet. Digitaalisen maailman muutos vaatii uudenlaista ajattelua ja teknologiaosaamisemme hyödyntämistä uusilla tavoilla. Yksityisen ja julkisen sektorin tulee yhteistyössä luoda edellytyksiä yrittäjyyskosysteemeille, joissa aktiiviset yhteisöt voivat itse kehittää paikallisen digitaalisen talouden toimintamalleja. Erityisesti tulee tukea tiedepohjaisia startup -yrityksiä.

Kaila Urpo
CSC-Tieteen tietotekniikan keskus Oy