

LAUSUNTO

Eduskunnan liikenne- ja viestintävaliokunnalle

Asia: U 28/2021 vp Valtioneuvoston kirjelmä eduskunnalle komission ehdotuksesta Euroopan parlamentin ja neuvoston asetukseksi tekoälyn harmonisoiduksi sääntelyksi (Artificial Intelligence Act)

Electronic Frontier Finland - Effi ry kiittää mahdollisuudesta tulla kuulluksi tässä tärkeässä asiassa. Effi on vuonna 2001 perustettu kansalais- ja asiantuntijajärjestö, joka puolustaa perus- ja ihmisoikeuksia digitaalisessa toimintaympäristössä.

Effi suhtautuu myönteisesti tekoälyn sääntelyyn. On hyvä että Euroopan unioni on asiassa liikkeellä ja ilahduttavaa on että esityksestä käy ilmi että liikkeellä ollaan tosissaan.

Tekoälyn määritelmässä on ongelmia

Esityksen määritelmät sisältävä artikla 3 määrittelee tekoälyn seuraavasti:

Tekoälyjärjestelmällä tarkoitetaan ohjelmistoa, joka on kehitetty yhdellä tai useammalla liitteessä I luetellulla tekniikalla ja lähestymistavalla ja joka voi tuottaa tiettyjä ihmisen määrittelemiä tavoitteita varten tuotoksia, kuten sisältöä, ennusteita, suosituksia tai päätöksiä. (käännös valtioneuvoston kirjelmästä)

Liitteessä I mainitaan seuraavat tekniikat tai teknologiat (käännökset allekirjoittaneen):

(a) Koneoppimismenetelmät sisältäen valvottu tai ohjattu (supervised) oppiminen, valvomaton tai ohjaamaton (unsupervised) oppiminen ja vahvistusoppiminen sisältäen laajan valikoima eri menetelmiä kuten syväoppiminen;

(b) Loogisiin sääntöihin ja tietoon perustuvat menetelmät sisältäen tiedon representoinnin, induktiivisen (loogisen) ohjelmoinnin, tietoarkistot, päättelyn ja deduktiiviset koneet, (symbolisen) päättelyn ja asiantuntijajärjestelmät;

(c) Tilastolliset menetelmät, bayesilainen arviointi, haku ja optimointimenetelmät.

Määritelmästä seuraa kaksi hankaluutta:

1) Moni nykyään laajalti käytetty menetelmä ja toimintatapa kuuluisi jatkossa säänneltävään alaan. Tämä on selkeintä tilastollisten menetelmien kohdalla. Jos tarkoitus on se, että jo olemassa olevia menetelmiä ei säänneltäisi tällä asetuksella, miten määritetty raja sen suhteen, milloin on jo kyse uudesta järjestelmästä? Jos esimerkiksi vanhaan ohjelmistoon tuodaan käsiteltäväksi dataa, jota siinä ei aiemmin ollut, pitäisikö sääntelyn koskea vain uuden datan käsittelyä vai kokonaisuutta vanha mukaan lukien? Entä jos vanhaan ohjelmistoon tuodaan uusi menetelmä: onko silloin kyse uudesta ohjelmistosta joka kuuluu sääntelyn alaan?

2) Tekoälyn käsitteen sitominen ohjelmistoon ei tee oikeutta tekoälyn olennaisimmalle piirteelle: datasta oppimiselle. Pelkkä ohjelmisto ilman sille syötettyä dataa ja siitä tehtyä oppimista on vasta alku eikä taatusti täytä tekoälyjärjestelmän tunnusmerkistöä - se on vain ohjelmisto.

Esimerkiksi jo nyt yleiset hermoverkko-sovellukset kykenevät oppimaan datasta itse. Tätä käytetään esimerkiksi konenäkösovelluksissa. Sääntelyn kannalta olennainen teoreettinen ongelma on se, että tätä "oppimista" ei pystytä kuvaamaan luonnollisella eli ihmisten käyttämällä kielellä. Se ei ole kuvattavissa sääntöperustaisesti. Konenäön tapauksessa voidaan nähdä, että kone tunnistaa tietyn esineen ja kykenee erottamaan sen toisista, mutta kone ei pysty kertomaan miten se sen tekee. Tätä ei myöskään pysty kertomaan ohjelmoija: hän pystyy vain sanomaan että tunnistaminen perustuu ohjelman muodostaman hermoverkon tilaan. Tämä lähestyy vanhaa kaskua kun perusteluita kysyttäessä vastaus kuuluu: "no se on siellä tietokoneella".

Esitys painottuu nähdäksemme liiaksi käsittelemään jo nyt käytössä olevia sääntöperustaisia tietojenkäsittelytapoja ja -menetelmiä. Sen pitäisi palvella paremmin tulevaisuuden tarpeita ja sääntelyn pitäisi olla toimivaa erityisesti itseoppivien järjestelmien kohdalla. Erityisesti ajettavaa ohjelmakoodia autonomisesti kehittävästä järjestelmästä tulisi mainita. Esimerkiksi järjestelmän, joka kykenee kirjoittamaan autonomisesti suoritettavaa ohjelmakoodia ja joka kykenee tuottamaan itsenäisesti toimivia "jälkeläisiä", tulisi itsestään selvästi olla sääntelyn piirissä.

Sääntelyn kannalta on yksi teoreettinen erityinen hankaluus. Mitä enemmän mennään itseoppiviin järjestelmiin ja kohti autonomista tekoälyä, sitä vähemmän tekoälyjärjestelmän tuotokset ja päätelmät ovat kuvattavissa ja ilmaistavissa sääntöperustaisesti tai luonnollisella kielellä. Vaikka esityksessä asetetaan velvoite tallentaa tekoälyjärjestelmän lokeja, mitä merkitystä niillä on, jos lokien kertomaa ei pystytä ilmaisemaan ihmisten käyttämällä kielellä? Läpinäkyvyysvaatimuksia tekoälyjärjestelmän toiminnasta tulisi konkretisoida ja vahvistaa.

Biometrinen tietojen käsittelystä

Erilaisten biometrinen tietojen käyttö tekoälysovelluksissa sisältää vaaroja. Tekoälysovellus käsittelee sille syötettyä dataa, kuten sormenjälkiä tai kasvokuvia, tietyn algoritmin mukaan. Syötteestä muodostetaan tietynlainen pseudotunniste. Esimerkiksi passeissa oleva, sormenjäljestä muodostettu pseudotunniste on joukko tietyistä kohdista otettuja, tietyllä tavalla käsiteltyjä pistemäisiä tunnusmerkkejä. Se, miten tämä on tarkkaan ottaen tehty ei ole turvallisuussyistä julkista tietoa.

Jos edellä mainittu tieto olisi julkista olisi täysin mahdollista, että joku tekisi keinotekoisien sormenjäljen käyttämällä pseudotunnistetta lähtötietona. Takavuosina kansanedustaja Kasvi ihmetteli että vaihdetaanko ihmisiltä sormet jos passirekisterin sormenjälkitietokanta sattuu vuotamaan. Nyt riskiskenaariot ovat toisenlaisia. Esimerkiksi kasvokuvien kohdalla teknologia on jo nyt niin pitkällä, että on mahdollista luoda synteettisiä kasvoja. Jos tiedossa on myös se, mitä piirteitä tunnistusalgoritmi etsii kasvoista, on mahdollista yrityksen ja erehdyksen menetelmällä luoda synteettisiä kasvoja, joilla on todellisen luonnollisen henkilön tunnusmerkit ja jotka kelpaavat kyseiselle tunnistusalgoritmille.

Riski saattaa tuntua tieteisfantasialta mutta kun kasvojentunnistukseen tarkoitettuja tuotteita on markkinoilla hyvinkin rajallinen määrä, riskin toteutuminen on täysin mahdollista. Kännyköiden kasvontunnistukseen perustuvan lukituksen huiputtaminen onnistuu jo nyt varsin helposti ja tekoälyllä pystytään luomaan sormenjälkiä, jotka kelpaavat sormenjälkien tunnistusjärjestelmille.

Kasvojentunnistusohjelmistojen kohdalla on toinenkin riski: jos tietty tekoälyjärjestelmä tekee jonkun kasvoista pseudotunnisteen, se on teknisesti siirrettävissä toiseen samalla ohjelmistolla toimivaan tekoälyjärjestelmään. On hyvin helppo keksiä tapoja käyttää tällaista mahdollisuutta väärin.

Sotilasteknologian ja muun käytön raja on häilyvä

Asetusta ei sovellettaisi tekoälyjärjestelmiin, jotka on kehitetty tai joita käytetään yksinomaan sotilastarkoituksiin.

Mitä tämä muotoilu tarkoittaisi autonomisten aseiden kaltaisten tekoälyä hyödyntävien laitteiden tapauksessa? Voisiko esimerkiksi lainvalvoja käyttää esimerkiksi käyttää räjähdettä tai lamauttavaa kaasupanosta kuljettavaa, autonomisesti toimivaa, tietyn henkilön lamauttamiseen tai surmaamiseen ohjelmoitua tekoälyä hyödyntävää dronea voimankäytön välineenä? Entä saisiko vastaavia välineitä kehittää ja myydä vaikka metsästyskäyttöön, esimerkiksi tietynlaisten yksilöiden metsästämiseen riistaeläinpopulaatiosta? Miten ja millä keinoin valvottaisiin sitä, ettei johonkin sallittuun tarkoitukseen kehitettyä laitetta käytetä väärään tarkoitukseen?

Asetusluonnoksen tavoite on pitää yksinomaan sotilaskäyttöön kehitetyt ja käytettävät tekoälyjärjestelmät tämän asetuksen sääntelyn ulkopuolella. Autonomisten aseiden sääntely, mielellään kieltäminen, on tehtävä kansainvälisellä tasolla. Ei tule elätellä lapsenuskoa siihen etteivätkö sotilasteknologian uudet tekoälyinnovaatiot leviäisi muihinkin käyttöihin. Jos näin kävisi, esitys ei anna tilanteeseen mitään muuta ohjenuoraa kuin sen itsestäänselvyyden että silloin alunperin sotilaskäyttöön kehitettyä teknologiaa ei käyttäisi enää vain sotilaskäytössä ja se olisi siten sääntelyn alaista. Näiden teknologoiden vaarallisuuden vuoksi sääntelyyn tulisi kuitenkin kiinnittää erityistä huomiota.

Suomen kieli ja muut pienet kielet ja tekoälypohjainen kieliteknologia

Jos tekoälyn määritelmä pysyy esitetyn laajuisena, myös useimmat kieliteknologian sovellukset tulevat sääntelyn piiriin. Internetissä toimivan teollisuuden suhtautuminen pieniin kielialueisiin on ollut näissä perin lepsu. Yksi ajankohtainen esimerkki on Youtuben kyvyttömyys tunnistaa koronavirukseen liittyvät videot yrityksen lupaamalla tavalla. Tällaiset lapsukset johtuvat tyypillisesti siitä, että oletetaan että muitakin kieliä kuin englantia voi käsitellä samoilla kieliteknologisilla menetelmillä. Erityisesti morfologisesti

rikkaan ja syntaktisesti monimutkaisen suomen kielen kohdalla moni englannin kielessä toimiva menetelmä ei yksinkertaisesti toimi. Monesti esimerkiksi sanojen esiintymistiheyteen ja esiintymisfrekvenssiin pohjautuvat menetelmät epäonnistuvat jo siksi, että taivutettuja suomen sanoja ei kyetä palauttamaan perusmuotoon analyysiä varten.

Vaikka emme ole varmoja siitä, onko kieliteknologiasovellusten lukeminen kokonaisuudessaan tekoälysäätelyn alaan järkevää, näemme tässä mahdollisuuden. Pienen kielialueen maan on mahdollista haastaa erilaisia palveluntarjoajia huonon ja joko yksilöllistä tai jopa yleistä haittaa aiheuttavan kieliteknologian käytöstä. Tätä mahdollisuutta muidenkaan pienempien kielialueiden ei tule väheksyä.

Elias Aarnio
varapuheenjohtaja
Electronic Frontier Finland – Effi ry