FCAI

Finnish Center for Artificial Intelligence

Petri Myllymäki
Professori (tekoäly ja koneoppiminen), Helsingin yliopisto
Varajohtaja, Suomen tekoälykeskus FCAI

Lausunto perustuslakivaliokunnalle 29.09.2021

Aalto University

UNIVERSITY OF HELSINKI

VTT

Ehdotuksesta tekoälyn sääntelyyn

- *KOM(2021) 202 lopullinen / valtioneuvoston kirjelmä U 28/2021 vp*
- Asetusehdotuksen olennaisena, **kannatettavana tavoitteena** on luoda Euroopasta ihmiskeskeisen tekoälyn kehittämiselle ja soveltamiselle suotuisa teknologisesti edistyksellinen ja innovaatiopositiivinen, mutta samaan aikaan ihmisoikeuksia ja yhteisiä arvojamme kunnioittava, turvallinen toimintaympäristö.
- Tähän tavoitteeseen pääsemiseksi on ensiarvoisen tärkeää, että tekoälyn sääntelyä koskeva asetus on mahdollisimman selkeä, yksiselitteinen, ristiriidaton, toimenpiteiltään oikein mitoitettu ja hyvin perusteltu.
- **Asetuksen nykyisessä muodossa tämä ei kaikilta osiltaan toteudu, vaan tavoitetilan saavuttamiseksi on ehdotusta vielä parannettava**.

Perusongelma

- Oikeudellisen kehyksen luomista lähestytään toisaalta
  a) riskiperustaisesti ja käyttötapauskohtaisesti, ja tämän rinnalla toisaalta
  b) teknologiariippuvasti, kohteena tekoälyjärjestelmät
- **Tämä dualistinen lähestymistapa ei toimi**, ja luo runsaasti epäkohtia ja epäselvyyksiä
- Tilanne voidaan korjata helposti siten, että asetusehdotus muutetaan teknologiariippumattomaksi, niin että se keskittyy ainoastaan käyttötapauksiin ja niihin liittyvien riskien torjumiseen **teknologisista ratkaisuista riippumatta**
  - asetus siis koskisi (kielletyissä tai korkean riskin käyttötarkoituksissa) **kaikkia tietokonejärjestelmiä**, eikä ainoastaan tekoälyjärjestelmiä
  - Koska asetus siis koskisi kaikkia järjestelmiä, myös tekoälyjärjestelmiä, niin tämä ei tekisi asetuksesta ”heikompaa”, vaan olennaisilla kohdealueilla päinvastoin vahvemman
- Teknologiariippumattomuus tekisi asetuksesta **selkeän ja ymmärrettävän ja aikaa kestävän**
  - kaikki tulevaisuudessa kehitettävät uudetkin teknologiat kuuluisivat automaattisesti sen piiriin, mikäli käyttötarkoitus kuuluu

Ongelma I: ei-tekoälyjärjestelmät kielletyissä käyttötarkoituksissa

- Asetusehdotus kieltää joukon digitaalisia tuotteita tai palveluja silloin kun niiden toteutus perustuu tekoälyyn
- Sitomalla käyttötarkoitukset tekoälyjärjestelmien määritelmään, asetus luo samalla luokan “ei-tekoälyjärjestelmät”, jolloin syntyy vaikutelma siitä, että **kielletyt käyttötarkoitukset ovatkin sallittuja kunhan käytetty teknologia ei täytä tekoälyjärjestelmän määritelmää** – tämä ei liene tarkoitus?
- Huomattakoon että kiellettyihin käyttötarkoituksiin on usein helppo tehdä hyvin yksinkertaisiakin teknologisia ratkaisuja, jotka eivät muistuta tekoälyjärjestelmiä
  - esimerkkinä vaikkapa sosiaalinen pisteytys syntymämaan mukaan
- Mikäli komission kanta on, että ei-tekoälyjärjestelmiä koskevaa uutta asetusta ei tarvita, koska nykyiset lait jo estävät kielletyt käyttötarkoitukset, herää kysymys että miksi näin ei olisi myös tekoälyjärjestelmien kanssa?
  - Päällekkäiset lait eivät lisää selkeyttä, päinvastoin

Ongelma II: ei-tekoälyjärjestelmät korkean riskin käyttötarkoituksissa

- Korkean tiskin käyttötarkoituksiin rakennettaviin järjestelmiin liitetään joukko vaatimuksia, joiden tarkoitus on taata että rakennettavat järjestelmät eivät aiheuta kohtuuttomia riskejä tai merkittäviä haittoja
  - Tämä on hyvä tavoite, kunhan huolehditaan siitä että vaatimukset eivät aiheuta järjestelmien kehittäjille sellaista haittaa joka antaa muille mantereille kilpailuetua
- Vaatimukset ovat suurelta osin **teknologiariippumattomia** prosessien laadunvalvontaan, testaukseen ja dokumentointiin liittyviä hyviä käytäntöjä
  - miksi näitä vaatimuksia ei kohdisteta korkean riskin käyttötarkoituksissa kaikkiin järjestelmiin vaan ainoastaan tiettyyn osajoukkoon tietokonejärjestelmiä (eli tekoälyjärjestelmiin)?

Ongelma III: tekoälyjärjestelmiä ei (voida) määritellä

- Tekoälyä ei asetusehdotuksessa oikeasti määritellä, vaan annetaan lähinnä (asiantuntijan silmin sekavasti rakennettu) lista erilaisia tekoälytutkimuksen osa-alueita, joita puolestaan ei määritellä lainkaan: **tämä ei ole sellainen määritelmä jonka perusteella voi ratkaista onko joku järjestelmä tekoälyä vai ei**.
  - Asetuksen antama oikeus komissiolle laventaa listaa myöhemmin ei poista tätä ongelmaa
- Tätä ongelmaa ei voi ratkaista, koska tekoäly on laaja tutkimusalue, joka muodostuu lukuisista heterogeenisistä osa-alueista ja niihin liittyvistä erilaisista teknologioista, jotka kehittyvät jatkuvasti, ja **on mahdotonta määritellä tekoälyjärjestelmien ja ei-tekoälyjärjestelmien välistä rajaa selkeästi**
  - Luo runsaasti kiistanalaisia tapauksia ja keinoja kiertää sääntelyä
- Lisäksi tekoäly ei koskaan muodosta yksin mitään tietokonejärjestelmää vaan on aina osa kokonaisuutta
  - kuinka suuren osan kokonaisuudesta pitää tekoälykomponenttien muodostaa jotta kokonaisuus tulkitaan tekoälyjärjestelmiksi?

Ongelma IV: yliharmonisointi

- Asetuksen oikeudellinen perusta (SEUT 114) **monopolisoi sääntelyn EU:ssa asetuksen kohdealueella**, mutta asetuksen duaaliluonteen vuoksi on epäselvää mikä tämä kohdealue on:
  - Kielletyt ja korkean riskin käyttötarkoitukset, vai tekoälyteknologiat, vai molemmat?
- Mikäli tekoälyjärjestelmät on asetuksen kohde, tästä seurannee se, että yksikään jäsenmaa ei saa jatkossa enää säännellä tekoälyä lainkaan, **käyttötarkoituksesta riippumatta**
  - tämä ns. ”pre-emption” EI SIIS rajoitu vain kiellettyihin ja korkean riskin tapauksiin, mikäli tekoäly on asetuksen kohde
- Koska asetuksen yritys määritellä tekoälyä on hyvin lavea, se kattaa suuren joukon menetelmiä, sisältäen mm. sääntöpohjaiset järjestelmät
  - Esimerkiksi lausuntokierroksella oleva laki hallinnon automatisoidusta päätöksenteosta Suomessa kuuluu selvästi EU:n asetuksen piiriin sen nykymuodossa, eli olisi jatkossa luultavasti kielletty
- Jäsenmaiden olisi myös mahdotonta luoda regulaatiota joka suosisi esimerkiksi vähäenergisiä, vihreää siirtymää tukevia tekoälyratkaisuja
- Määritelmässä mainitaan myös logiikkaan perustuvat järjestelmät, ja tämä on mahdollista tulkita jopa niin että kaikki tietokonejärjestelmät kuuluvat määritelmän piiriin (koska tietokonekielet perustuvat logiikkaan): koskeeko monopolisointi itse asiassa kaikkia digitaalisia järjestelmiä?

Yhteenveto

- Asetusehdotuksessa linkitetään yhteen säänneltävät riskianalyysiin perustuvat käyttötarkoitukset ja joukko tekoälyksi kutsuttavia teknologiota, mutta tätä teknologiaperhettä on mahdotonta määritellä sääntelyyn soveltuvalla tavalla, mikä tekee asetuksesta **epäselvän ja vaikeasti sovellettavan ja luo joukon vakavia ongelmia**, **toimien hyviä tarkoitusperiään vastaan**.
- **Tekoälyn määritelmää ei kuitenkaan onneksi edes tarvita** vaan siitä voidaan hyvin luopua ja tehdä asetuksesta **teknologiariippumaton**, mikä tekee asetuksesta selkeän ja ymmärrettävän ja aikaa kestävän, koska kaikki tulevaisuudessa kehitettävät uudetkin teknologiat kuuluvat silloin automaattisesti sen piiriin silloin kun käyttötarkoitus kuuluu.
- Komission voi PR-syistä olla enää vaikeaa luopua termistä ”tekoälyasetus”, mutta asia voidaan korjata asetuksen otsikkoa muuttamatta esimerkiksi joko
  a) toteamalla että asetuksessa tekoälyjärjestelmäksi kutsutaan mitä tahansa tietokonejärjestelmää jota voi käyttää kielletyssä tai korkean riskin käyttötarkoituksessa,
  b) tai toteamalla erikseen että ”tekoälyasetus” on teknologiariippumaton, koskien tekoälyjärjestelmiä ja muita järjestelmiä yhtä lailla, mutta otsikossa esiintyy sana ”tekoäly” koska suuri osa asetuksen kohdealueella toimivista järjestelmistä saattaa käyttää tekoälyä jossain muodossa.