Turun yliopisto, Tekoälyakatemia

Lausunto

31.03.2026

Asia: VN/12980/2025

Luonnos hallituksen esitykseksi digipalvelulain 6 a §:n muuttamisesta (tekoäly neuvonnassa)

Lausunnonantajan lausunto

Voitte kirjoittaa lausuntonne alla olevaan tekstikenttään

Keskusteluagenttien käyttö osana palveluja on perusteltua ja mahdollistaa kohdennetun neuvonnan. Generatiivisiin kielimalleihin perustuvat järjestelmät pystyvät käymään monimutkaisia keskusteluja ja antamaan hyödyllisiä tietoja, ainakin jos kysymys on riittävän yksinkertainen ja siihen liittyviä esimerkkitapauksia löytyy järjestelmän koulutusaineistosta.

Generatiivisten kielimallien riskinä on "hallusinointi", asiaan kuulumattomien, väärien tai harhaanjohtavien vastausten antaminen. Mallien kehittyessä niiden luotettavuus paranee, mutta hallusinointi on tämänkaltaisten tekoälymallien perustavanlaatuinen ominaisuus, eikä niiden voida odottaa toimivan aina täysin johdonmukaisesti tai luotettavasti. Tästä huolimatta kielimalliteknologia voi tarjota paljon apua kohdennetussa tiedonhaussa, ja sen käyttöönotto voi tuoda etuja moniin palveluihin, kunhan myös sen riskit ja rajoitteet tunnistetaan.

Lakiesitykseen liittyen nostaisimme esille neljä seikkaa, joihin tulisi kiinnittää erityistä huomiota:

1) Kielimallien yksityisyys ja tietoturva on huomioitava. Tekoälyjärjestelmän käyttö ei saa tarkoittaa, että asiakkaan tietoja siirretään vähemmän tietoturvalliseen ympäristöön kuin missä niitä muuten käytettäisiin. Parhaiten tämä varmistuisi käyttämällä vain lokaaleja kielimalleja, ja välttämällä ulkomaisia pilvipalveluita. Tekoälyteknologia kehittyy nopeasti ja edistyneimpienkin mallien ominaisuudet päätyvät pian kilpaileviin toteutuksiin, joten vaihtoehtoja (myös lokaalisti käytettäviä) on paljon tarjolla. Esimerkiksi Turun yliopiston TurkuNLP-tutkimusryhmä on kehittänyt avoimia kielimalleja (Poro yms.) erityisesti suomenkielisten käyttäjien tarpeisiin.

2) Suurin hyöty toiminnan tehostamisessa saavutettaisiin automatisoimalla päätöksentekoa, mutta tähän liittyvät myös suurimmat riskit. Jos kuitenkin on tilanteita, joissa päätös on ilmeinen saatavilla olevien tietojen valossa (esim. jonkin tuen myöntäminen) voisi olla mahdollista, että automaatiolla tuotettua päätöstä tarjottaisiin asiakkaalle. Tällöin on tärkeää, että asiakas tietää päätöksen perustuvan automaation, ja voi tarvittaessa aina pyytää ihmistä arvioimaan sen uudelleen.

3) Käyttäjän on aina tiedettävä, asioiko koneen vai ihmisen kanssa, kuten EUn tekoälylainsäädäntökin edellyttää. On vaikea nähdä miten tämä tieto voisi olla niin "ilmeistä", että siitä ei tarvitsisi käyttäjää erikseen informoida. Luotettava informointi ei kuitenkaan vaadi sen enempää kuin selvää merkintää asiakaspalvelukeskustelijan käyttäjätunnuksessa siitä, onko kyseessä kone vai ihminen. Tämä pitäisi olla hyvin toteutettavissa missä tahansa kyseeseen tulevassa järjestelmässä. Voisi myös miettiä, olisiko mahdollista kehittää laajempaa standardointia (esim. yleisesti käytettyjä symboleita) kertomaan, ollaanko keskustelemassa ihmisen vai tekoälyn kanssa.

4) Tekoälyjärjestelmien riskinä on hallusinaatio ja harhaanjohtavat vastaukset. Näitä riskejä ei voi selvittää kattavasti etukäteen, ja vaikka tekoälyjärjestelmä toimisi yleensä luotettavasti, siinä on aina virhevastausten mahdollisuus. Niinpä järjestelmän luotettavuuden varmistamiseksi sitä tulee seurata käytännön neuvontatilanteissa jatkuvasti ja pitkäjänteisesti, ja käytyjä keskusteluja on analysoitava järjestelmän toiminnan varmistamiseksi. Generoidut neuvontatilanteet on säilytettävä niin kauan kuin tarvitaan niiden seurausten arvioimiseksi. Tietoja tulee voida käyttää anonymisoidusti tutkimuskäyttöön järjestelmän toiminnan ja laadun varmistamiseksi. Käyttäjällä tulee olla mahdollisuus ladata kaikki käymänsä keskustelut, ja käyttäjän tulee voida pyytää henkilökohtaisten tietojensa poistoa.

Yllä olevat neljä kohtaa perustuvat yleisiin käytäntöihin tekoäly- ja kielimallijärjestelmien turvallisesta ja vastuullisesta käytöstä. Toivomme, että nämä seikat huomioidaan osana laissa jo mainittuja laadun varmistamisen, oikeusturvan ja syrjimättömyyden vaatimuksia, sekä osana lakiehdotuksen perusteluja. Voimassa olevan lain kohtien 1)-3) poistaminen on perusteltua tekoälyagenttien käyttöä ajatellen, mutta tällöin on myös huomioitava tekoälymallien riskit, joita edellä toimme esille.

Itse lakitekstiin toivoisimme muutosta ehdotuksen kohtaan 1), eli käyttäjälle on aina kerrottava yksiselitteisesti asioiko hän ihmisen vai koneen kanssa. Nämä seikat huomioiden suosittelemme lakiehdotuksen täsmentämistä seuraavaan muotoon:

Muokattu ehdotus:

6 a § Palveluautomaation käyttö neuvonnassa

Viranomainen voi antaa hallinnon asiakkaalle digitaalisessa palvelussa neuvontaa, joka perustuu reaaliaikaiseen viestien vaihtoon palvelun käyttäjän ja palveluautomaation välillä, jos palvelun käyttäjälle:

1) kerrotaan, että hän vaihtaa viestejä palveluautomaation kanssa;

2) kerrotaan, miten hän voi ottaa yhteyttä viranomaiseen asioinnin jatkamiseksi luonnollisen henkilön kanssa; ja

3) tarjotaan mahdollisuus tallentaa palveluautomaation kanssa käyty viestien vaihto.

Viranomaisen on varmistettava neuvonnan laatu ennen palveluautomaation käyttöönottoa ja valvottava neuvonnan laatua käytön aikana. Viranomaisen on hallittava palveluautomaation käytöstään palvelun käyttäjien oikeusturvaan ja syrjimättömyyteen kohdistuvia riskejä.

Viranomaisen on nimettävä 2 momentin mukaisten velvoitteidensa toteuttamisesta vastuussa oleva henkilö. Viranomaisen palveluksessa olevaan julkisyhteisön työntekijään sovelletaan hänen tässä pykälässä tarkoitettuja tehtäviä hoitaessaan virkamiehen rikosoikeudellista virkavastuuta koskevia säännöksiä, viraltapanoseuraamusta lukuun ottamatta.

Björne Jari
Turun yliopisto, Tekoälyakatemia