Tekoäly vastuullisuusnäkökulmasta

Viime aikoina tekoäly on saanut paljon huomiota julkisuudessa mm., Italian tietosuojaviranomaisen määrättyä väliaikaisen kiellon italialaisten henkilötietojen käsittelylle Chat GPT-tekoälyjärjestelmässä. Kiellon syynä oli erityisesti henkilötietojen käsittelyn läpinäkymättömyys ja muut tietosuojapuutteet. Eettiseltä kulmalta tarkasteltuna päätöksessä korostui lasten asema tekoälyn sidosryhminä, sillä kyseisen järjestelmän ei katsottu riittävällä tavalla kykenevän huomioimaan sisällön soveliaisuutta lapsille. Organisaatioiden tuleekin varmistua, että tekoäly ei pelkästään täytä juridisia vaatimuksia, vaan tekoälyn hyödyntämisen riskejä sidosryhmille sekä sen eettisyyttä arvioidaan huolellisesti, jotta tekoälyjärjestelmä kestää kriittistä tarkastelua eri toimijoiden kuten viranomaisten, sijoittajien, työntekijöiden, asiakkaiden ja laajemmin yhteiskunnan puolelta.

Organisaatiot hyödyntävät lisääntyvissä määrin erilaisia tekoälyjärjestelmiä aina teknisten toimintojen tehostamisesta työntekijöiden ohjaamiseen ja asiakasprosessin sujuvoittamiseen sekä myynnin kehittämiseen. Tekoälyjärjestelmiä voidaan hyödyntää laajasti hyvin erilaisissa prosesseissa, jolloin niihin ei välttämättä kohdistu runsaasti sidosryhmiin liittyviä juridisia reunaehtoja kuten tietosuoja- tai kuluttajansuojasääntelyä. Toisaalta tekoälyä hyödynnettäessä usein käytännössä käsitellään tavalla tai toisella henkilötietoja esimerkiksi tekoälyn opetusvaiheessa, jolloin tulee huomioida sääntelyn vaatimukset. Tekoälyn kohdistuessa luonnollisiin henkilöihin, organisaation tulee myös tekoälyjärjestelmän tilaajana ja rekisterinpitäjänä varmistua, että hankittava järjestelmä täyttää esimerkiksi tietosuoja-asetuksen vaatimukset.

EU:ssa on valmisteilla tekoälyä ja dataa koskevaa sääntelyä, mutta jo nykyisellään tietosuoja-asetus (GDPR) asettaa luontevan kehikon, jonka kautta tekoälyn vaatimustenmukaisuutta on mahdollista tarkastella myös henkilötietojen käsittelyä laajemmin. Erityisesti GDPR:n sisältämien tietosuojaperiaatteiden kautta voidaan tunnistaa vaatimuksia tekoälyjärjestelmälle. Eräitä olennaisia periaatteita ovat läpinäkyvyyden periaate, jonka perusteella henkilötietoja tulee käsitellä rekisteröidyn kannalta läpinäkyvästi. Tekoälyn kontekstissa haasteita läpinäkyvyydelle aiheutuu sen monimutkaisesta päättelyketjusta ja prosesseista. Voidaan myös pohtia sitä, ymmärrämmekö todella sitä, millaisia arvovalintoja tekoälyn kehittämiseen ja sen toimintaan sisältyy: miten tekoäly ymmärtää esimerkiksi sitä, onko sen tekemä päätelmä tai suositus moraalisesti hyväksyttävä vallitsevissa olosuhteissa? Ratkaisut tulee löytää ja niitä tulee pohtia etupainotteisesti. Esimerkkeinä voidaan nostaa esille huolellinen suunnittelu (esim. opetusdatan laadun varmistaminen), riskiarviointi ja ihmisten suorittama jatkuva valvonta.

Tekoälyn opettamiseen tarvitaan tyypillisesti suurta määrää dataa, joka usein sisältää jonkinlaisia henkilötietoja. GDPR edellyttää, että tietoja käsitellään vain, mikäli käsittelylle voidaan tunnistaa oikeusperuste. Tekoälyn kontekstissa olennaista on pohtia sitä, millä perusteella esimerkiksi työntekijöiden tai asiakkaiden henkilötietoja on vastuullista ja laillista käyttää tekoälyn opettamiseen.

Tekoälyjärjestelmä tuottaa teknisestä toteutustavasta riippuen suuren määrän päätelmiä ja olettamuksia ihmisistä, jolloin GDPR:n tietojen täsmällisyyden vaatimus korostuu. Tällöin olennaista on varmistua siitä, että tiedoista käy selvästi ilmi, että ne eivät ole faktatietoa (esimerkiksi luovutettaessa tietoa eteenpäin tai käytettäessä tietoa muihin tarkoituksiin). Tässä yhteydessä myös yksilöiden oikeuksia saada vaikuttaa omien tietojensa käsittelyyn tulee tarkastella. Mitkä ovat sidosryhmien oikeudet vaikuttaa siihen, mitä tietoja ja mihin tarkoituksiin tekoäly hyödyntää? GDPR:n sisältämät rekisteröityjen oikeudet tarjoavat tähän luontevan lähestymistavan.

Entä mitä organisaatioiden tulee tehdä, jotta ne voivat varmistua siitä, että tekoälyn hyödyntäminen on juridisten reunaehtojen mukaista minimoiden siitä aiheutuvat riskit sidosryhmille sekä eettisen tarkastelun kestävää?

Seuraavat askelmerkit auttavat edellä kuvattujen kokonaisuuksien haltuunotossa:

- Määritetään juridiset reunaehdot ja varmistetaan niiden noudattaminen: Tekoälyn käyttökohteet vaikuttavat soveltuvaan sääntelyyn (esimerkiksi sidosryhmät, hyödynnettävä data, päätöksenteon vaihe, automaation tekninen toteutustapa) ja sitä kautta sallittaviin toteutustapoihin ja tarvittaviin kontrolleihin.
- Laaditaan tekoälyn riskienarviointi: Kun tekoälyn konteksti ja vaatimukset on tunnistettu, laaditaan tekoälyn riskienarviointi, jossa arvioidaan tekoälyn sidosryhmille aiheuttamia potentiaalisia riskejä ja tunnistetaan teknisiä, organisatorisia ja sopimuksellisia suojatoimia riskien vähentämiseksi hyväksyttävälle tasolle.
- Tekoälyn eettisyyden varmistaminen: Juridisten reunaehtojen noudattamisen ja riskien vähentämisen lisäksi tulee varmistua, että tekoälyn hyödyntäminen on eettistä. Tässä auttaa esimerkiksi tekoälyn eettisten periaatteiden laatiminen ja eettisyyden huomioiminen osana koko tekoälyn elinkaarta.

Tekoälyn vaatimustenmukaisuus, riskienhallinta ja eettisyys edellyttävät kokonaisuuksien tarkastelua yhdessä ja erikseen. Tärkeää on, että organisaatiolla on ymmärrys tekoälyn teknisestä toiminnasta ja tietovirroista sekä rooleista ja vastuista tekoälyratkaisun toimittajan suuntaan. Vastuuta ei ole mahdollista ulkoistaa toimittajalle, sillä tekoälyä hyödyntävällä organisaatiolla on vastuu omista sidosryhmistään ja käsiteltävistä henkilötiedoista. Toimittajienhallinta on kuitenkin olennainen osa edellä kuvattujen kokonaisuuksien tehokasta jalkauttamista. Lopulta vastuu tekoälyn eettisyydestä on jokaisen sitä hyödyntävän organisaation harteilla. Vastuullinen tekoäly tarkoittaa myös sitä, että organisaatiolla on rohkeutta olla tinkimättä eettisistä periaatteistaan myös tekoälyn hyödyntämisen osalta.

Jere Lehtioksa, lakimies
Jeren työtehtävänä on tarjota KPMG:n asiakkaille tukea teknologiaregulaatiota kuten dataa, tietosuojaa ja tekoälyä koskevissa muutos-, riskienhallinta- ja kehitysprojekteissa.

Jere on osallistunut muun muassa datan jakamisen sopimusehtojen laatimista ja tietosuojan reunaehtojen arviointia ja riskienhallintaa koskeviin projekteihin. Jerellä on kokemusta myös tekoälyn riskienhallinnasta ja regulaatiosta mukaan lukien automaattisen päätöksenteon lainmukaisuuden arvioinneista ja laajoista riskienarviointiprojekteista. Jerellä on vahva käytännön osaaminen myös datan ja tekoälyn teknisestä puolesta, ja hän on osallistunut useiden projektien toteuttamiseen yhdessä KPMG:n teknisten asiantuntijoiden kanssa.