U 28/2021 vp Valtioneuvoston kirjelmä eduskunnalle komission ehdotuksesta Euroopan parlamentin ja neuvoston asetukseksi tekoälyn harmonisoiduksi sääntelyksi (Artificial Intelligence Act), asiassa saatu valtioneuvoston jatkokirjelmä 8.9.2023

Tässä lausunnossa tarkastellaan Euroopan parlamentin muutosehdotusta, jolla generatiivisen tekoälyn tarjoajille asetettaisiin velvollisuus julkaista yhteenveto mallin kouluttamiseen käytetystä tekijänoikeudella suojatusta aineistosta. Kyseisen kohdan (muutosehdotuksen 28 b artikla) mukaan:

> ”4. Tekoälyjärjestelmissä, jotka on tarkoitettu tuottamaan eriasteisen itsenäisesti sisältöä, kuten kehittynyttä teksti-, kuva-, ääni- tai videosisältöä, jäljempänä ’generatiivinen tekoäly’, käytettävien perusmallien tarjoajien ja tarjoajien, jotka ovat erikoistuneet perustusmallien integrointiin generatiiviseen tekoälyjärjestelmään, on lisäksi
>
> a) noudatettava 52 artiklan 1 kohdassa säädettyjä läpinäkyvyysvelvoitteita;
>
> b) koulutettava perusmalli ja tarvittaessa suunniteltava se ja kehitettävä sitä siten, että varmistetaan riittävät suojatoimet unionin lainsäädännön vastaisen sisällön tuottamisen estämiseksi yleisesti tunnustetun alan viimeisimmän kehityksen mukaisesti ja rajoittamatta perusoikeuksia, sananvapaus mukaan lukien;
>
> c) *dokumentoitava ja asetettava julkisesti saataville riittävän yksityiskohtainen yhteenveto tekijänoikeuslainsäädännön nojalla suojatun koulutusdatan käytöstä, sanotun kuitenkaan rajoittamatta unionin tai kansallisen tekijänoikeuslainsäädännön soveltamista.*” (korostus lisätty)

Generatiivisen tekoälyn kehittämisessä yleensä tapahtuu tekijän yksinoikeuksien piiriin kuuluvaa kappaleiden valmistamista koulutusaineistoon kuuluvista suojan kohteista (kopiointia), jos esimerkiksi koneoppimismallin kouluttamisessa käytetään tekijänoikeudella suojattua aineistoa. Toiminta siten joko edellyttää oikeudenhaltijoiden suostumusta tai että jokin tekijänoikeuden rajoitus soveltuu toimintaan. Myös generatiiviseen tekoälyyn perustuvaa palvelua tarjottaessa saattaa tapahtua yksinoikeuksien piiriin kuuluvaa toimintaa (mm. koulutusaineiston oikeuksia loukkaavaa yleisön saataviin saattamista).

Generatiivisen tekoälyn osalta ehdotettu velvollisuus julkaista yhteenveto kouluttamiseen käytetystä aineistosta edistäisi tekijänoikeuksien täytäntöönpanoa, lisensointia ja hallinnointia. Oikeudenhaltijoiden on tällä hetkellä vaikeaa päätellä tarjotuista palveluista tai muilla keinoilla saada tietoa siitä, onko niiden suojattua aineistoa käytetty generatiivisen tekoälyn kouluttamiseen. Velvollisuus myös kannustaisi

generatiivisen tekoälyn kehittäjiä ottamaan tekijänoikeudelliset kysymykset huomioon jo tekoälykehityksen toteuttamista suunnitteluvaiheessa.

U-jatkokirjeessä (5.9.2023) ilmaistut Suomen kannat siitä, että 1) velvollisuuden toteuttamisen tulisi olla käytännössä mahdollista eikä kohtuuttoman raskasta ja että 2) velvollisuuden suhdetta ja vaikutuksia tekijänoikeusääntelyyn tulisi arvioida tarkemmin, ovat kuitenkin perusteltuja. Alla esitetään eräitä näihin näkökohtiin liittyviä huomioita.

Velvollisuuden sisällön oikeasuhtaisuus

Jotta velvollisuus olisi käytännössä toteutettavissa eikä kohtuuttoman raskas täytettäväksi, yhteenvedon sisältöä koskevien vaatimusten tulisi vastata tekoälykehityksen olosuhteita ja vaatimuksia. Koska tekoälyn kehittämisessä käytettävät aineistot voivat olla hyvin laajoja, yksityiskohtaisia tietoja ei ole realistista edellyttää. Sen sijaan esimerkiksi keskeiset lähteet, joista aineisto on hankittu, ja keinot, joilla data on kerätty, voisivat realistisesti olla generatiivisen tekoälyn kehittäjän ilmoitettavissa. Tekoälyn kehittäjän tulee esimerkiksi joka tapauksessa varmistaa, että käytettäviin aineistoihin on laillinen pääsy, jos kehittäjä nojautuu tekstin ja datanlouhintaa koskevaan tekijänoikeuden poikkeukseen (DSM-direktiivin 3 tai 4 artikla). Myös kun koulutusaineisto on hankittu sopimuksen perusteella, esimerkiksi lisensoimalla, lähde on jo kehittäjän tiedossa.

Velvollisuuden sisällön tulisi olla oikeasuhtainen. Sitä koskevien yksityiskohtaisempien vaatimusten tulisi olla tavoiteltavien (tekijänoikeudellisten) intressien edistämisen kannalta perusteltuja, eikä vaatimusten tulisi rajoittaa muiden intressien toteutumista, kuten tekoälyn kehittämisen edellytyksiä, enempää kuin on suhteellisuusperiaatteen mukaista. Velvollisuuden sisällöstä ja soveltamisesta tulisi olla tarjolla suuntaviivoja ja ohjeistusta, jonka laatimisessa otetaan huomioon sekä tekoälyn kehittämis- ja hyödyntämisolosuhteet että tekijänoikeudelliset näkökohdat.

Muutosehdotuksessa on epäselvää, mitä vaatimuksia dokumentointia koskeva maininta c-alakohdassa asettaa mahdollisesti yhteenvedon julkaisemista koskevan velvollisuuden lisäksi. Myös dokumentointia mahdollisesti koskevan velvollisuuden tulisi sisällöltään olla suhteellisuusperiaatteen mukainen ja sisällöltään olosuhteet huomioon ottava, ja sitäkin koskien tulisi olla saatavilla suuntaviivoja ja muuta ohjeistusta.

Velvollisuuden vaikutukset ja suhde tekijänoikeuslainsäädäntöön

Tekijänoikeuden ja yhteenvetoa koskevan velvollisuuden välisistä yhteyksistä tulee huomata varsinkin velvollisuuden suhde tekstin ja datanlouhintapoikkeukseen (DSM-direktiivin 3 ja 4 artikla). Yleinen tekstin ja datanlouhintapoikkeus (DSM-direktiivin 4 artikla) mahdollistaa oikeudenhaltijalle louhintaoikeuden pidättämisen, jonka seurauksena poikkeus ei sovellu ja tekoälykehittäjä ei voi siihen nojautua. Yhteenvetoa koskevan velvollisuuden tarkemman sisällön osalta tulisi pyrkiä ratkaisuihin, jotka

tukevat tämän pidättämiseen liittyvän mekanismin toimivuutta niin tekijänoikeuden haltijoiden kuin tekoälykehittäjienkin kannalta tarkasteltuna.

Velvollisuuden soveltamisala

Velvollisuus koskisi sisällöntuotantoon tarkoitettuja perustamalleja (generatiivinen tekoäly). Perustamallilla tarkoitetaan ”tekoälyjärjestelmämallia, joka on koulutettu skaalattavasti laajamittaisella datalla, suunniteltu tuottamaan yleisluonteisia tuotoksia ja jota voidaan mukauttaa monenlaisiin eri tehtäviin” (Euroopan parlamentin muutosehdotus 3 artiklan 1 kohdan 1 c alakohdaksi).

Velvollisuus ei soveltuisi niihin sisällöntuotantoon tarkoitettuihin tekoälyjärjestelmiin, jotka eivät samalla ole ehdotetun määritelmän mukaisia perustamalleja. On epäselvää, onko velvollisuus kohdennettu tarkoituksenmukaisesti, koska sen ulkopuolelle jäisi sellaisia sisällöntuotantoon tarkoitettuja tekoälyjärjestelmiä, joihin liittyvät tekijänoikeudelliset kysymykset voivat olla yhtä merkittäviä kuin niiden, joita velvollisuus koskee. Esimerkiksi perustamallin määritelmän edellyttämä mukautuvuus tai tuotosten yleisluonteisuus saattaa puuttua joidenkin nimenomaan sisällöntuotantoon kehitettyjen tekoälysovellusten osalta.

Syväväärennöksiä koskeva läpinäkyvyysvelvoite ja tekijänoikeudet

Euroopan parlamentti on lisäksi ehdottanut niin sanottuja syväväärennöksiä koskevien läpinäkyvyysvelvoitteiden osalta, että jos ”sisältö on osa selvästi luovaa, satiirista, taiteellista tai fiktiivistä elokuvaa, videopelien kuvamateriaalia ja vastaavaa teosta tai ohjelmaa, 3 kohdassa säädetyt läpinäkyvyysvelvoitteet rajoittuvat siihen, että tällaisesta tuotetusta tai manipuloidusta sisällöstä ilmoitetaan asianmukaisesti, selkeästi ja näkyvästi tavalla, joka ei haittaa teoksen esittämistä, ja että tapauksen mukaan ilmoitetaan sovellettavat tekijänoikeudet.” (muutosehdotuksen 52 artiklan 3 kohdan 1 alakohta). Käytetty ilmaisu ”sovellettavat tekijänoikeudet” (”the applicable copyrights, where relevant”) on monitulkintainen ja epäselvä, koska siitä ei voi päätellä mihin tekijänoikeuksiin kohdassa viitataan ja mitä sanalla ”sovellettavat” tarkoitetaan.

18.9.2023

IPR University Center

Juha Vesala, johtaja