

Avoimuusvelvoitteet

Webinaari: Tekoälyasetuksen ajankohtaiskatsaus 8.10.
Mari Rupponen, juristi



AI Act

Avoimuusvelvoitteiden tavoitteet

- ▶ Avoimuusvelvoitteet edellyttävät toimijoilta tekoälyn käytöstä kertomista sekä tekoälyn tuottaman tai muokkaaman sisällön merkitsemistä tietyissä tilanteissa.
- ▶ Velvoitteiden keskeisenä tavoitteena on varmistaa, että ihmiset tietävät milloin he ovat tekemisissä tekoälyn kanssa ja voivat arvioida sen tuottamaa sisältöä asianmukaisesti.



Käsitteet: Tarjoaja ja käyttöönottaja

Tarjoaja

- ▶ Luonnollinen tai oikeushenkilö, viranomainen, virasto tai muu elin, joka
- ▶ kehittää tai kehityttää tekoälyjärjestelmän tai yleiskäyttöisen tekoälymallin ja,
- ▶ saattaa sen markkinoille tai ottaa tekoälyjärjestelmän käyttöön omalla nimellään tai tavaramerkkillään.

Käyttöönottaja

- ▶ Luonnollinen tai oikeushenkilö, viranomainen, virasto tai muu taho, joka
- ▶ käyttää valvonnassaan olevaa tekoälyjärjestelmää
- ▶ muuhun kuin yksityiseen käyttöön.

Avoimuusvelvoitteet

50(1): Luonnollisten henkilöiden kanssa suoraan vuorovaikutuksessa oleva järjestelmä.

• **Tarjoajien** on suunniteltava järjestelmä siten, että luonnolliset henkilöt tietävät olevansa tekemisissä tekoälyjärjestelmän kanssa.

50(2): Järjestelmä tuottaa tai manipuloi synteettisiä kuvia, videoita, ääntä tai tekstiä.

• **Tarjoajien** on varmistettava, että järjestelmä merkitsee tuotokset koneluottavassa muodossa ja että ne ovat tunnistettavissa tekoälyjärjestelmän tuottamiksi.

50(3): Tunteentunnistaminen tai biometrinen kategorisointi.

• **Käyttöönottajien** tulee ilmoittaa luonnollisille henkilöille, että he ovat edellä mainitun kohteena.

50(4): Järjestelmä tuottaa tai manipuloi syvävääreännöksiä tai on julkaistu teksti, joka koskee yleistä etua.

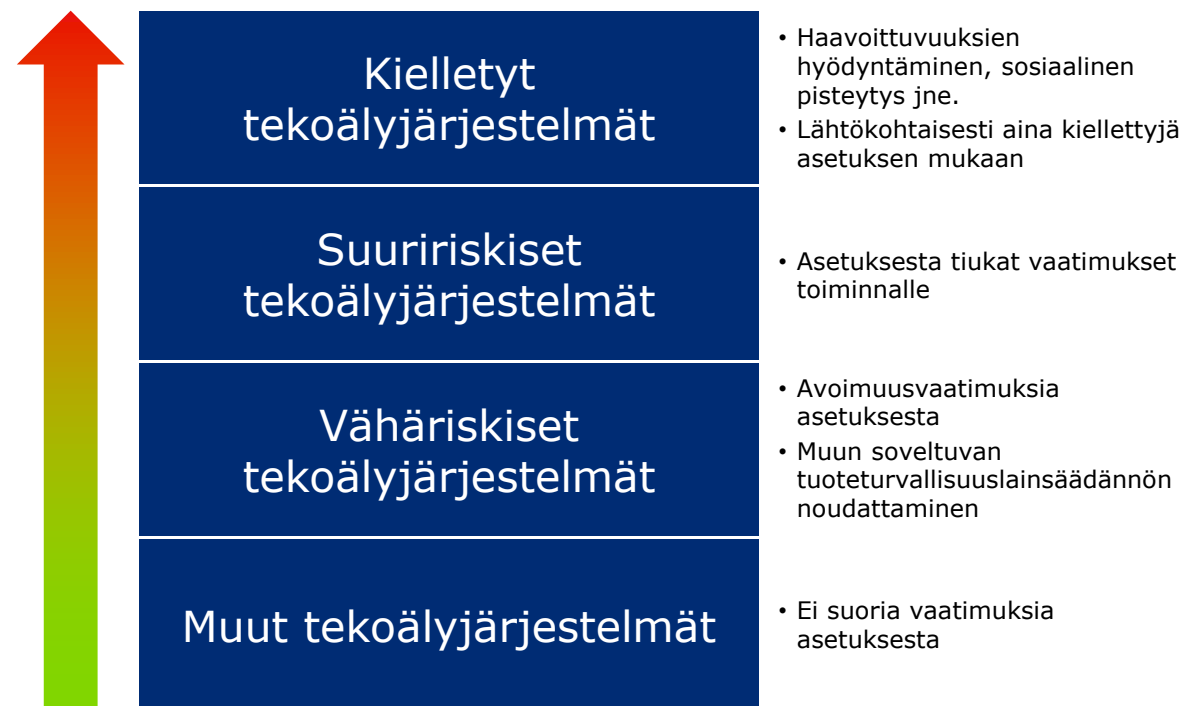
• **Käyttöönottajien** tulee merkitä, että sisältä on tekoälyn tuottamaa tai manipuloitua.

Artiklan 1–4 kohdassa tarkoitetut tiedot on annettava luonnollisille henkilöille selkeällä ja erottuvalla tavalla viimeistään ensimmäisen vuorovaikutuksen tai altistumisen yhteydessä (art. 50(5)).

Tietojen on lisäksi täytettävä sovellettavat saavutettavuusvaatimukset.

Avoimuusvelvoitteiden soveltamisala

- ▶ Sama järjestelmä voi kuulua sekä **avoimuusvelvoitteiden** että **suuririskisiä järjestelmiä** koskevan sääntelyn soveltamisalaan.
 - ▶ Velvoitteet ovat tällöin rinnakkaisia.
- ▶ Artiklan 50 velvoitteiden noudattaminen ei tarkoita, että tekoälyjärjestelmän käyttö olisi sallittua.
 - ▶ Järjestelmä voi silti kuulua kiellettyjen käytäntöjen piiriin.



Artikla 50(1)

Luonnollisten henkilöiden kanssa suoraan
vuorovaikutuksessa oleva järjestelmä



Luonnollisten henkilöiden kanssa suoraan vuorovaikutuksessa olevat järjestelmät

Elementti	Kuvaus	Keskeinen sisältö	Rajaukset
1. AI-järjestelmä	Järjestelmän tulee täyttää AI-järjestelmän määritelmä (art. 3(1)).	Edellyttää varsinaista tekoälyä; pelkkä automaatio ei riitä.	Ei kata esim. perinteisiä poissaoloviestejä tai sääntöpohjaisia pikavastauksia.
2. Tarkoitettu vuorovaikutukseen	Järjestelmä on suunniteltu vuorovaikutukseen ihmisten kanssa.	Kaksisuuntainen tiedon tai toiminnan vaihto; useita muotoja (teksti, ääni, visuaalinen, fyysinen).	Ei kata pelkkää passiivista datankeruuta (esim. kasvojentunnistus kulunvalvonnassa).
3. Suora vuorovaikutus	Vuorovaikutus tapahtuu suoraan AI-järjestelmän ja henkilön välillä.	Tyypillisesti reaaliaikaista tai lähes reaaliaikaista; järjestelmä tuottaa vastauksen tai toiminnan suoraan henkilölle.	Ei kata välillistä käyttöä (esim. asiakaspalvelija käyttää AI-työkalua taustalla).
4. Luonnollisten henkilöiden kanssa	Vuorovaikutus kohdistuu luonnollisiin henkilöihin.	Voi koskea käyttäjiä, ammattilaisia tai muita henkilöitä.	Ei kata järjestelmiä, jotka toimivat suljetuissa fyysisissä ympäristöissä (esim. teollisuusympäristö).

Vuorovaikutuksesta ilmoittaminen

- ▶ Tieto annettava selkeästi viimeistään ensimmäisen kontaktin yhteydessä ja tarvittaessa vuorovaikutuksen aikana.
 - ▶ Esimerkiksi chatbotin ensimmäisen viestin yhteydessä.
- ▶ Ilmoituksen tulee yksiselitteisesti ilmaista vuorovaikutuksen keinotekoinen (ei-inhimillinen) luonne.
 - ▶ Esimerkiksi "Olet vuorovaikutuksessa tekoälyjärjestelmän kanssa". Viestiä voi tehostaa esimerkiksi kuvakkeella.
- ▶ Ilmoitusta on mukautettava kohderyhmän mukaan.
 - ▶ Huomioitava esimerkiksi lapset ja ikäihmiset.



Olen tekoälyavustaja. En anna hoito-ohjeita tai terveysneuvontaa – terveysasioissa ole yhteydessä terveydenhuollon ammattilaiseen. Vastaukseni perustuvat sivustomme sisältöön ja voivat sisältää virheitä. Antamani hinta-arviot ovat alustavia ja voivat muuttua. Tarkistathan lopullisen hinnan ajanvarauksen yhteydessä.



Hei, olen tekoälyllä toimiva avustajasi ja autan sinua löytämään tietoa ja asioimaan sujuvasti. Ethän jaa minulle henkilö- tai terveystietojasi. Miten voin olla avuksi?

Keskivertokäyttäjä

- ▶ Ilmoitusta ei tarvita, jos keskivertokäyttäjä ymmärtää selvästi asioivansa tekoälyn kanssa.
- ▶ Keskivertokäyttäjällä tarkoitetaan kohtuullisen hyvin informoitua, tarkkaavaista ja huolellista henkilöä.
- ▶ Arvioon vaikuttavat:
 - ▶ Käyttötilanne ja järjestelmän kohderyhmä ml. haavoittuvat ryhmät.
 - ▶ Vuorovaikutuksen ilmeisyys.
- ▶ Tulkitaan kapeasti.



Artikla 50(2)

Tekoälyjärjestelmä, joka tuottaa tai manipuloi synteettisiä kuvia, videoita, ääntä tai tekstiä



Tekoälyjärjestelmä, joka tuottaa tai manipuloi synteettisiä kuvia, videoita, ääntä tai tekstiä

- ▶ Artikla 50(2) velvoittaa **tarjoajia** varmistamaan, että tekoälyjärjestelmän tuotokset merkitään koneellisesti luettavassa muodossa ja että ne tunnistettavissa tekoälyjärjestelmän tuottamiksi tai manipuloimiksi.
- ▶ Luomisen tai muokkaamiseen tulee ylittää ns. vakiomuokkaamisen raja.
 - ▶ Avustava vakiomuokkaus
 - ▶ Poikkeus koskee järjestelmiä, jotka tekevät vain tavanomaista muokkausta esim. kielioppi, muotoilu, tekninen laatu.
 - ▶ Ei-merkittävä muutos syötteeseen
 - ▶ Poikkeus soveltuu, jos tekoäly ei olennaisesti muuta syötteen sisältöä tai merkitystä esim. oikoluku, kuvan kevyt rajaus tai melunpoisto.



Soveltamisalaan kuuluvat sisällöt

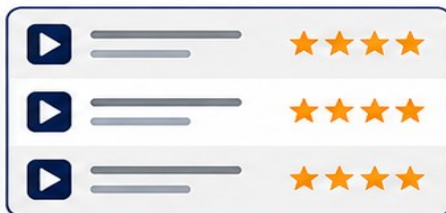
Teksti	Kuva	Ääni	Video	Multimodaalinen sisältö
				
Erillinen symbolinen sisältö (merkit tai numerot), joka on ihmisten luettavissa ja tulkittavissa.	Staattinen visuaalinen esitys, joka välittää tietoa yhdestä ajanhetkestä ja joka on ihmisten nähtävissä.	Ajassa vaihteleva signaali, joka sisältää ääntä (esim. puhe, musiikki tai muut äänisignaalit) ja on ihmisen kuuloaistin havaittavissa.	Kuvasarja, joka voi olla synkronoitu äänen kanssa ja on ihmisten nähtävissä.	Sisältö, joka koostuu synteettisen sisällön eri muotojen yhdistelmästä.

Artiklan 50(2) soveltamisalaan kuuluvat myös:

Virtuaali-, lisätty tai yhdistetty todellisuus	Tekoälyn tuottamat tai muokkaamat 3D-kuvat, videot ja äänet	Digitaaliset kaksoset	Toiminnalliset (agenttiset) tekoälyjärjestelmät
			
Koska virtuaali- ja lisätty todellisuus muodostuvat ajassa etenevästä kuvasekvenssistä, niitä voidaan pitää "videosisältönä".		Virtuaaliset kopiot henkilöistä, fyysisistä kohteista tai järjestelmistä.	Jos tekoälyagentti toteuttaa toiminnon, joka on luonnollisten henkilöiden havaittavissa äänen, kuvan, videon tai tekstin muodossa, tuotokset kuuluvat soveltamisalaan.

Soveltamisalan ulkopuolelle jäävät sisällöt

1. Vain olemassa olevan sisällön jäljentäminen, valinta tai järjestäminen



Esimerkiksi musiikkisoihtolistat tai suositusjärjestelmät, jotka valitsevat tai järjestävät olemassa olevaa sisältöä käyttäjän mieltymysten tai toiminnan perusteella ilman uuden sisällön luomista tai muokkaamista.

2. Havainto-, tallennus- ja tiedonsiirtosisältö ilman muutoksia.



Esimerkiksi anturien, kameroiden, robottien tai älymittareiden tallenteet; verkon jännite- ja taajuusmittaukset; ajoneuvojen GPS-tiedot; tiedonsiirrot ilman muutoksia.

3. Tuotokset, joita ei ole tarkoitettu ihmisen havaittavaksi



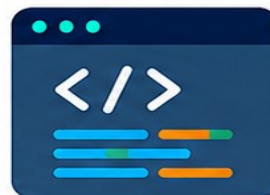
Tuotokset, jotka välittyvät koneelta koneelle ja joita käsitellään automaattisesti.

4. Muut esimerkit soveltamisalan ulkopuolelta

Lyhyet numerosarjat,
merkit tai kirjaimet

7 3 8 5 2
% @ \$
A B 7 K 9
...

Lähdekoodi



Suljetut teolliset tai
tuotekehitysympäristöt



Merkintä- ja tunnistamisvelvoite

- ▶ Artikla 50(2) asettaa tekoälyjärjestelmien **tarjoajille** kaksiosaisen velvoitteen:
 - I. Järjestelmän tuotosten tulee olla merkitty koneluettavassa muodossa, ja
 - II. Niiden tulee olla tunnistettavissa tekoälyn tuottamiksi tai manipuloimiksi.
- ▶ Molempien vaatimusten on täytyttävä samanaikaisesti - pelkkä merkintä ilman tunnistettavuutta ei täytä artiklan vaatimuksia.
- ▶ Merkintä- ja tunnistusratkaisujen tulee olla tehokkaita, luotettavia, kestäviä ja yhteentoimivia.



Artikla 50(3)

Tunteentunnistaminen tai biometrinen
kategorisointi



Tunteentunnistaminen ja biometrinen luokittelu

- ▶ Artikla 50(3) velvoittaa järjestelmien **käyttöönottajat** ilmoittamaan luonnollisille henkilöille, että heihin kohdistuu tunteentunnistus- tai biometrinen luokittelujärjestelmä.
 - ▶ Tunteentunnistusjärjestelmä: AI-järjestelmä, joka tunnistaa tai päättelee henkilön tunteita tai aikomuksia biometrisista tiedoista.
 - ▶ Biometrinen luokittelujärjestelmä: AI-järjestelmä, joka luokittelee henkilöitä biometrisen datan perusteella tiettyihin kategorioihin.
- ▶ Henkilölle tulee ilmoittaa selkeästi, että häneen kohdistuu tunteentunnistus tai biometrinen luokittelu.



Artikla 50(4)

Tekoälyjärjestelmä, joka tuottaa tai manipuloi syväväärennöksiä tai on julkaistu teksti, joka koskee yleistä intressiä



Tekoälyjärjestelmä, joka tuottaa tai manipuloi syväväärennöksiä

- ▶ Artikla 50 (4.1) velvoittaa **käyttöönottajia** ilmoittamaan, kun tekoälyjärjestelmällä on tuotettu tai manipuloitu syväväärennöksen muodostavaa kuva-, ääni- tai videosisältöä.
- ▶ Pienet tekniset muokkaukset eivät yleensä tee sisällöstä syväväärennöstä (esim. värikorjaus, taustan hienosäätö, kohinanpoisto).
 - ▶ Arvio on kuitenkin tapauskohtainen ja riippuu siitä, vaikuttaako muutos aidon vaikutelman syntymiseen.
- ▶ Arvioitaessa tuotetun sisällön luonnetta ei käytetä "keskivertokäyttäjän" kriteeriä, vaan huomioidaan yleisön koostumus.
 - ▶ Erityisen tärkeää tämä on, jos yleisöön kuuluu lapsia, ikääntyneitä tai henkilöitä, joilla on heikommat digitaaidot.



Syväväärennöksen käsite

Samankaltaisuus (resemblance)

Sisällön tulee muistuttaa kohdettaan merkittävästi. Täydellinen vastaavuus ei ole tarpeen, mutta samankaltaisuuden arviointi tehdään tapauskohtaisesti.

Olemassaoleva (existing)

Sisällön tulee muistuttaa sellaista kohdetta, joka on tai voisi olla olemassa todellisuudessa. Epärealistinen sisältö (esim. luonnonlakien vastainen) ei kuulu määritelmän piiriin.

Syväväärennös, art. 3(60)

Virheellinen aitouden vaikutelma

Sisällön tulee johtaa harhaan tai erehdyttää aitoudesta tai totuudenmukaisuudesta. Arviointi koskee sitä, voiko sisältö harhauttaa käyttäjää – ei järjestelmän käyttäjän tarkoitusta.

Henkilöt, esineet, paikat, toimijat tai tapahtumat

Realistiset ihmiset, konkreettiset objektit, todelliset sijainnit, elävät olennot, realistiset tilanteet tai tapahtumakokonaisuudet.

Esimerkki: syvävääreännöksen merkitseminen

Merkinä



Ei merkinä



Taiteellinen, luova, humoristinen tai fiktiivinen sisältö

- ▶ Tietyissä tilanteissa riittää asianmukainen ilmoitus, joka ei häiritse teoksen esittämistä tai kokemista. Tämä mahdollisuus on olemassa, kun:
 - ▶ Syväväärennös-sisältö (kuva, ääni tai video) on osa selvästi taiteellista, luovaa, satiirista, fiktiivistä tai analogista teosta tai ohjelmaa.
 - ▶ On ilmeisestä, että sisältö kuuluu johonkin näistä kategorioista.
- ▶ Arviointi tehdään tapauskohtaisesti.



Teksti, joka koskee yleistä intressiä

- ▶ Artikla 50(4.2) velvoittaa **käyttöönottajia** ilmoittamaan, kun tekoälyjärjestelmää on käytetty tekstin tuottamiseen tai manipuloimiseen, ja sisältö julkaistaan tarkoituksena informoida yleisöä yleistä etua koskevissa asioissa.
- ▶ Tekoälyn tuottaman tai manipuloiman tekstin tulee olla julkaistu eli saatavilla määrittelemättömälle ja riittävän suurelle yleisölle.
 - ▶ Esimerkiksi verkkosivut, erilaiset maksulliset sisällöt.
- ▶ Tekstin tarkoituksena tulee olla tiedon, mielipiteiden tai faktojen välittäminen yleisölle.
- ▶ Sisällön tulee käsitellä yhteiskunnallisesti merkityksellisiä aiheita.
 - ▶ Esimerkiksi politiikka, julkishallinto, oikeudenhoito, perusoikeudet, turvallisuus, kansanterveys, ympäristö, kuluttajansuoja.



Inhimillinen tarkastus tai toimituksellinen valvonta

- ▶ Artikla 50(4) 2 alakohdan mukaisesta tiedonantovelvollisuudesta on poikkeus, jos kaksi ehtoa täyttyvät samanaikaisesti:
 - ▶ AI-tuotettu tai muokattu teksti on käynyt läpi inhimillisen tarkastuksen tai toimituksellisen valvonnan; ja
 - ▶ Julkaisusta vastaa luonnollinen henkilö tai oikeushenkilö, jolla on toimituksellinen vastuu.
- ▶ Toimituksellinen valvonta tarkoittaa vastuullisen toimijan tosiasiallista valtaa hyväksyä, muuttaa tai hylätä sisältö substanssiperustein.
- ▶ Julkaisijalla tulee olla lopullinen oikeudellinen vastuu sisällöstä.
 - ▶ Vastuutahon (henkilö tai organisaatio) identiteetin ja yhteystietojen tulee olla julkisesti saatavilla.



EU:n viralliset tekoälykuvakkeet

AI

Tekoälyä hyödynnetty

Kun tekoälyä on käytetty tekstin tai syvävääreännös-materiaalin (kuva, video, ääni) tekemisessä.

Esimerkiksi video, jossa äänet on luotu tekoälyllä.

AI MODIFIED

Tekoälyllä muokattu

Kun aiemmin luotua, ihmisen tekemää sisältöä on osittain muokattu tekoälyllä.

Esimerkiksi, kun aidon valokuvan kasvot vaihdetaan poliitikon kasvoihin tekoälyllä tai aito valokuva tyhjästä asunnosta sisustetaan tekoälyn avulla.

AI GENERATED

Tekoälyllä luotu

Kun materiaali (kuva, ääni, video, teksti) on täysin tekoälyn tekemää.*

Esimerkiksi syvävääreännös-videot tai automaattiset uutistiivistelmät.

*Taiteellisiin, luoviin, fiktiivisiin tai satiirisiin teoksiin voi soveltua rajoitettuja velvollisuuksia.

Avoimuusvelvoitteiden valvonta



Liikenne- ja viestintävirasto Traficom

- ▶ Liikenne- ja viestintävirasto Traficom valvoo markkinavalvontaviranomaisena tekoälyasetuksen 50 artiklan mukaisten avoimuusvelvoitteiden noudattamista.
- ▶ Valvontaviranomaisen käytettävissä on kansallisen markkinavalvontalain mukaiset valvonnan toimenpiteet.
 - ▶ Viranomaiselle on mahdollisuus määrätä myös seuraamusmaksu.
- ▶ Viranomaisen toiminnassa painotetaan tällä hetkellä ohjeistusta ja neuvontaa.
- ▶ Havaitsemistaan 50 artiklan rikkeistä voi ilmoittaa viraston verkkosivuilta löytyvällä lomakkeella.



Muut valvontaviranomaiset

- ▶ Kansalliset markkina- ja valvontaviranomaiset
 - ▶ Suuririskisen tekoälyjärjestelmän tarjoajaan ja käyttöönottajaan kohdistuvien tekoälyasetuksen 50 artiklan mukaisten avoimuusvelvoitteiden noudattamista valvoo kuitenkin 3 §:n mukainen toimivaltainen markkina- ja valvontaviranomainen.
- ▶ Komissio
 - ▶ Artikla 50(1) ja (2), kun tekoälyjärjestelmä perustuu yleiskäyttöiseen tekoälymalliin (GPAI) ja sama tarjoaja kehittää sekä mallin että järjestelmän.
 - ▶ Artikla 50(3) ja (4) vain, jos sama tarjoaja, joka on kehittänyt sekä mallin että järjestelmän, on myös järjestelmän käyttöönottaja.



Lisätietoja

- ▶ Liikenne- ja viestintävirasto Traficomien [verkkosivut](#)
- ▶ Tekoälyllä tuotetun sisällön merkitsemistä koskevat EU:n [tekoälykuvakkeet](#)
- ▶ [Suuntaviivat](#) tiettyjen tekoälyjärjestelmien tarjoajien ja käyttöönottajien avoimuusvelvoitteista
- ▶ Tekoälyllä tuotetun sisällön avoimuutta koskevat [käytännäsäännöt](#)



Kiitos!