Tietosuojaneuvoston lausunto (64 artikla)

Lausunto 28/2024 tietyistä henkilötietojen käsittelyyn tekoälymallien yhteydessä liittyvistä tietosuojanäkökohdista.

Annettu 17. joulukuuta 2024

Translations proofread by EDPB Members.

This language version has not yet been proofread.

Tiivistelmä

Tekoälyteknologiat tarjoavat monia mahdollisuuksia ja hyötyjä useilla eri aloilla ja yhteiskunnallisissa toiminnoissa.

Tietosuojaan liittyvää perusoikeutta suojelemalla tietosuoja-asetus tukee näitä mahdollisuuksia ja edistää muita EU:n perusoikeuksia, kuten oikeutta ajatuksen, sanan ja tiedonvälityksen vapauteen, oikeutta koulutukseen ja elinkeinovapauteen. Siten yleinen tietosuoja-asetus on oikeudellinen kehys, joka kannustaa vastuulliseen innovointiin.

Ottaen huomioon näiden teknologioiden esiin tuomat tietosuojakysymykset Irlannin valvontaviranomainen pyysi tässä yhteydessä tietosuojaneuvostoa antamaan lausunnon yleisesti sovellettavista asioista yleisen tietosuoja-asetuksen 64 artiklan 2 kohdan mukaisesti. Pyyntö liittyy henkilötietojen käsittelyyn tekoälymallien kehittämis- ja käyttöönottovaiheissa. Pyynnössä kysyttiin seuraavaa: 1) milloin ja miten tekoälymallia voidaan pitää ”anonymisoituna”, 2) miten rekisterinpitäjät voivat osoittaa, että oikeutettu etu on asianmukainen oikeusperuste kehittely- ja 3) käyttöönottovaiheessa, ja 4) mitkä ovat seuraukset henkilötietojen laittomasta käsittelystä tekoälymallin kehittämisvaiheessa tekoälymallin käsittelylle tai toiminnalle myöhemmin.

Ensimmäisen kysymyksen osalta lausunnon mukaan toimivaltaisten valvontaviranomaisten olisi arvioitava väitteet tekoälymallin anonymiteetistä tapauskohtaisesti, sillä tietosuojaneuvosto katsoo, että henkilötietojen perusteella koulutettuja tekoälymalleja ei voida kaikissa tapauksissa pitää anonyymeinä. Jotta tekoälymallia voitaisiin pitää anonyyminä, 1) todennäköisyyden (myös probabilistisuus) sille, että mallia kehitettäessä henkilötietoja on poimittu suoraan henkilöistä , joiden henkilötietoja on käytetty mallin kehittämisessä, ja 2) todennäköisyyden sille, että tällaisia henkilötietoja saadaan tarkoituksellisesti tai tahattomasti kyselyistä, olisi oltava vähäinen ottaen huomioon *kaikki ne keinot, joita rekisterinpitäjä tai muu henkilö voi kohtuullisen todennäköisesti käyttää*.

Arviointia varten valvontaviranomaisten olisi tarkasteltava rekisterinpitäjän toimittamia asiakirjoja, joilla osoitetaan mallin anonymiteetti. Tältä osin lausunnossa on ohjeellinen ja ei-tyhjentävä luettelo menetelmistä, joita rekisterinpitäjät voivat käyttää anonymiteetin osoittamisessa. Siten valvontaviranomaiset voivat ottaa tämän huomioon arvioidessaan rekisterinpitäjän anonymiteettiä koskevaa vaatimusta. Tämä kattaa esimerkiksi lähestymistavat, joita rekisterinpitäjät ovat omaksuneet kehittämisvaiheessa estääkseen tai rajoittaakseen koulutukseen käytettävien henkilötietojen keräämistä, vähentääkseen tietojen tunnistettavuutta, estääkseen niiden poimimisen tai antaakseen varmuuden kyvystä kestää hyökkäyksiä, joissa hyödynnetään uusinta tekniikkaa.

Toisen ja kolmannen kysymyksen osalta lausunnossa esitetään yleisiä näkökohtia, jotka valvontaviranomaisten on otettava huomioon arvioidessaan, voivatko rekisterinpitäjät vedota oikeutettuun etuun asianmukaisena oikeusperusteena tekoälymallien kehittämisen ja käyttöönoton yhteydessä tapahtuvalle käsittelylle.

Lausunnossa muistutetaan, että yleisessä tietosuoja-asetuksessa säädettyjen oikeusperustojen välillä ei ole hierarkiaa ja että rekisterinpitäjien tehtävänä on määritellä asianmukainen oikeusperuste käsittelytoiminnalleen. Tämän jälkeen lausunnossa muistutetaan kolmivaiheisesta testistä, joka olisi suoritettava, kun arvioidaan oikeutetun edun käyttöä oikeusperusteena. Siihen kuuluu 1) rekisterinpitäjän tai kolmannen osapuolen tavoitteleman kyseisen oikeutetun edun yksilöinti, 2) käsittelyn tarpeellisuuden analysointi tavoitellun oikeutetun edun (etujen) kannalta (jäljempänä

tarpeellisuustesti) ja 3) sen arvioiminen, etteivät rekisteröityjen edut tai perusoikeudet ja -vapaudet syrjäytä oikeutettua etua (etuja) (jäljempänä tasapainotesti).

Ensimmäisen vaiheen osalta lausunnossa muistutetaan, että etu voidaan katsoa oikeutetuksi, jos seuraavat kolme kumulatiivista kriteeriä täyttyvät: etu on 1) lainmukainen, 2) ilmaistu selkeästi ja täsmällisesti ja on 3) todellinen ja läsnä (ei spekulatiivinen). Kiinnostus voi kohdistua esimerkiksi tekoälymallin kehittämiseen, kuten käyttäjiä avustavan keskustelevan tekijän palveluun tai sen käyttöönottoon liittyvään kehittämiseen, jotta parannettaisiin uhkien havaitsemista tietojärjestelmässä.

Toisen vaiheen osalta lausunnossa muistutetaan, että tarpeellisuuden arviointi edellyttää seuraavien seikkojen tarkastelua: 1) mahdollistaako käsittely oikeutetun edun tavoittelun ja 2) voiko tätä etua tavoitella vähemmän puuttuvalla tavalla. Arvioidessaan, täyttyykö välttämättömyyden edellytys, valvontaviranomaisten olisi kiinnitettävä erityistä huomiota käsiteltävien henkilötietojen määrään ja siihen, onko se oikeassa suhteessa tavoiteltavaan oikeutettuun etuun, myös tietojen minimointiperiaatteen valossa.

Kolmannen vaiheen osalta lausunnossa muistutetaan, että tasapainotesti olisi suoritettava ottaen huomioon kunkin tapauksen erityisolosuhteet. Sen jälkeen siinä esitetään yleiskatsaus seikoista, jotka valvontaviranomaiset voivat ottaa huomioon arvioidessaan, syrjäyttävätkö rekisteröityjen edut, perusoikeudet ja -vapaudet rekisterinpitäjän tai kolmannen osapuolen edut.

Kolmannessa vaiheessa lausunnossa tuodaan esiin perusoikeuksiin kohdistuvia erityisiä riskejä, joita voi syntyä joko tekoälymallien kehittämis- tai käyttöönottovaiheissa. Siinä selvennetään myös, että tekoälymallien kehittämis- ja käyttöönottovaiheen aikana tapahtuva henkilötietojen käsittely voi vaikuttaa rekisteröityihin eri tavoin, mikä voi olla myönteistä tai kielteistä. Arvioidakseen tällaisia vaikutuksia valvontaviranomaiset voivat tarkastella mallien käsittelemien tietojen luonnetta, käsittelyn kontekstia ja käsittelyn mahdollisia muita seurauksia.

Lausunnossa korostetaan lisäksi rekisteröityjen kohtuullisten odotusten merkitystä tasapainotestissä. Tämä voi olla tärkeää, koska tekoälymalleissa käytettävät teknologiat ovat monimutkaisia ja rekisteröityjen voi olla vaikea ymmärtää niiden mahdollisia käyttötarkoituksia ja niihin liittyviä erilaisia käsittelytoimia. Tältä osin sekä rekisteröidyille annetut tiedot että käsittelyn asiayhteys voivat kuulua niihin seikkoihin, jotka on otettava huomioon arvioitaessa, voivatko rekisteröidyt kohtuudella odottaa, että heidän henkilötietojaan käsitellään. Asiayhteydessä tämä voi sisältää seuraavat seikat: onko henkilötiedot julkisesti saatavilla vai ei, rekisteröidyn ja rekisterinpitäjän välisen suhteen luonne (ja se, onko näiden kahden välillä yhteys), palvelun luonne, yhteys, jossa henkilötiedot on kerätty, lähde, josta tiedot on kerätty (eli verkkosivusto tai palvelu, jossa henkilötiedot on kerätty, ja tarjotut yksityisyysasetukset), mallin mahdolliset muut käyttötavat ja se, ovatko rekisteröidyt tosiasiallisesti selvillä siitä, että heidän henkilötietojaan on verkossa.

Lausunnossa muistutetaan myös, että jos rekisteröityjen edut, oikeudet ja vapaudet näyttävät olevan rekisterinpitäjän tai kolmannen osapuolen tavoittelevien oikeutettujen etujen edellä, rekisterinpitäjä voi harkita lieventävien toimenpiteiden käyttöönottoa rajoittaakseen käsittelyn vaikutusta kyseisiin rekisteröityihin. Lieventäviä toimenpiteitä ei pidä sekoittaa toimenpiteisiin, jotka rekisterinpitäjän on lain mukaan joka tapauksessa toteutettava yleisen tietosuoja-asetuksen noudattamisen varmistamiseksi. Lisäksi toimenpiteet olisi räätälöitävä tapauksen olosuhteiden ja tekoälymallin ominaisuuksien, mukaan lukien sen käyttötarkoituksen, mukaisesti. Tältä osin lausunnossa esitetään ei-tyhjentävä luettelo esimerkkejä lieventävistä toimenpiteistä, jotka liittyvät kehittämisvaiheeseen (myös verkkosivujen haravoinnin osalta) ja käyttöönottovaiheeseen. Lieventävät toimenpiteet voivat

kehittyä nopeasti, ja ne olisi mukautettava tapauksen olosuhteisiin. Sen vuoksi valvontaviranomaisten tehtävänä on arvioida tapauskohtaisesti toteutettujen lieventävien toimenpiteiden asianmukaisuutta.

Neljännen kysymyksen osalta lausunnossa muistutetaan yleisesti, että valvontaviranomaisilla on harkintavalta arvioida mahdollista rikkomusta (rikkomuksia) ja valita asianmukaisia, tarpeellisia ja oikeasuhteisia toimenpiteitä ottaen huomioon kunkin yksittäisen tapauksen olosuhteet. Sen jälkeen lausunnossa tarkastellaan kolmea skenaariota.

Skenaariossa 1 henkilötietoja säilytetään tekoälymallissa (mikä tarkoittaa, että mallia ei voida pitää anonyyminä, kuten ensimmäisen kysymyksen tapauksessa) ja sama rekisterinpitäjä käsittelee henkilötietoja myöhemmin (esimerkiksi mallin käyttöönoton yhteydessä). Lausunnossa todetaan, että se, liittyykö kehitys- ja käyttöönottovaiheisiin eri tarkoituksia (ja ovatko ne siten eri käsittelytoimia) ja missä määrin alkuperäisen käsittelytoimen oikeusperusteen puuttuminen vaikuttaa myöhemmän käsittelyn laillisuuteen, olisi arvioitava tapauskohtaisesti tapauksen asiayhteydestä riippuen.

Skenaariossa 2 henkilötietoja säilytetään mallissa, ja toinen rekisterinpitäjä käsittelee niitä mallin käyttöönoton yhteydessä. Tältä osin lausunnossa todetaan, että valvontaviranomaisten olisi otettava huomioon, onko mallin käyttöön ottava rekisterinpitäjä suorittanut asianmukaisen arvioinnin osana vastuuvelvollisuuttaan osoittaakseen tietosuoja-asetuksen 5 artiklan 1 kohdan a alakohdan ja 6 artiklan, varmistaakseen, että tekoälymallia ei ole kehitetty käsittelemällä henkilötietoja lainvastaisesti. Tässä arvioinnissa olisi otettava huomioon esimerkiksi se, mistä henkilötiedot on saatu, ja mahdollinen rikkominen käsittelyn kehittämisvaiheessa, erityisesti jos valvontaviranomainen tai tuomioistuin olisi tullut siihen tulokseen. Arvioinnin yksityiskohtaisuus riippuu käsittelyn riskeistä käyttöönottovaiheessa.

Skenaariossa 3 rekisterinpitäjä käsittelee henkilötietoja laittomasti tekoälymallin kehittämiseksi ja varmistaa sen anonymisoinnin ennen kuin sama tai toinen rekisterinpitäjä aloittaa muun henkilötietojen käsittelyn käyttöönoton yhteydessä. Tältä osin lausunnossa todetaan, että jos voidaan osoittaa, että tekoälymallin myöhempi toiminta ei sisällä henkilötietojen käsittelyä, tietosuojaneuvosto katsoo, että yleistä tietosuoja-asetusta ei sovelleta. Näin ollen alkuperäisen käsittelyn lainvastaisuus ei saisi vaikuttaa mallin myöhempään toimintaan. Lisäksi tietosuojaneuvosto katsoo, että kun rekisterinpitäjät käsittelevät myöhemmin käyttöönottovaiheen aikana kerättyjä henkilötietoja sen jälkeen, kun malli on anonymisoitu, yleistä tietosuoja-asetusta sovellettaisiin näihin käsittelytoimiin. Näissä tapauksissa lausunnossa katsotaan, että yleisen tietosuoja-asetuksen osalta alkuperäisen käsittelyn lainvastaisuuden ei pitäisi vaikuttaa käyttöönottovaiheessa suoritetun käsittelyn lainmukaisuuteen.

Sisällys

ottaa huomioon luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta 27 päivänä huhtikuuta 2016 annetun Euroopan parlamentin ja neuvoston asetuksen (EU) 2016/679 (jäljempänä ”yleinen tietosuoja-asetus”) 63 artiklan ja 64 artiklan 2 kohdan,

ottaa huomioon ETA-sopimuksen sekä erityisesti sen liitteen XI ja pöytäkirjan 37, sellaisina kuin ne ovat muutettuina 6. päivänä heinäkuuta 2018 annetulla ETA:n sekakomitean päätöksellä N:o 154/2018[1],

ottaa huomioon työjärjestyksensä 10 ja 22 artiklan,

sekä katsoo seuraavaa:

1) Euroopan tietosuojaneuvoston (jäljempänä **tietosuojaneuvosto** tai **EDPB**) päätehtävänä on varmistaa yleisen tietosuoja-asetuksen (GDPR) yhdenmukainen soveltaminen Euroopan talousalueella (ETA). Yleisen tietosuoja-asetuksen 64 artiklan 2 kohdassa säädetään, että valvontaviranomainen, tietosuojaneuvoston puheenjohtaja tai komissio voi pyytää minkä tahansa yleisluonteisen tai useammassa kuin yhdessä jäsenvaltiossa vaikutuksia tuottavan asian käsittelyä tietosuojaneuvostossa lausunnon saamiseksi. Tämän lausunnon tarkoituksena on yleisluonteisen tai useammassa kuin yhdessä Euroopan talousalueen jäsenvaltiossa vaikutuksia tuottavan asian tutkiminen.

Tietosuojaneuvosto antaa lausunnon yleisen tietosuoja-asetuksen 64 artiklan 3 kohdan nojalla, yhdessä Euroopan tietosuojaneuvoston työjärjestyksen 10 artiklan 2 kohdan kanssa, kahdeksan viikon kuluessa ensimmäisestä arkipäivästä sen jälkeen, kun puheenjohtaja ja toimivaltaiset valvontaviranomaiset ovat katsoneet, että asiakirja on valmis. Puheenjohtajan päätöksellä tätä määräaikaa voidaan jatkaa kuudella viikolla käsiteltävän asian monimutkaisuus huomioon ottaen.

ON ANTANUT SEURAAVAN LAUSUNNON:

1 Johdanto

1.1 Tiivistelmä tosiseikoista

1. Irlannin valvontaviranomainen, jäljempänä **Irlannin valvontaviranomainen** tai **pyynnön esittävä valvontaviranomainen**, pyysi 4. syyskuuta 2024 tietosuojaneuvostoa antamaan yleisen tietosuoja-asetuksen 64 artiklan 2 kohdan mukaisesti lausunnon tekoälymalleista ja henkilötietojen käsittelystä, jäljempänä **pyyntö**.

2. Tietosuojaneuvoston puheenjohtaja ja Irlannin valvontaviranomainen katsoivat, että asiakirja-aineisto oli täydellinen 13. syyskuuta 2024. Tietosuojaneuvoston sihteeristö antoi tiedoston levitykseen seuraavana työpäivänä 16. syyskuuta 2024. Puheenjohtaja otti huomioon asian monimutkaisuuden ja päätti jatkaa lainmukaista määräaikaa yleisen tietosuoja-asetuksen 64 artiklan 3 kohdan ja Euroopan tietosuojaneuvoston työjärjestyksen 10 artiklan 4 kohdan mukaisesti.

[1] Viittauksilla ”jäsenvaltioihin” tarkoitetaan kauttaaltaan tässä lausunnossa Euroopan talousalueen jäsenvaltioita. Viittaukset ”unioniin” on tulkittava kauttaaltaan tässä lausunnossa viittauksina ETA:an.

3. Pyyntö koskee tiettyjä tekoälymallien koulutukseen, päivittämiseen, kehittämiseen ja käyttöön liittyviä tekijöitä, joissa henkilötiedot muodostavat osan asianomaista tietoaineistoa. Irlannin valvontaviranomainen korostaa, että pyyntö koskee keskeisiä kysymyksiä, joilla on suuri vaikutus rekisteröityihin ja rekisterinpitäjiin Euroopan talousalueella, eikä kansallisilla valvontaviranomaisilla ole tässä vaiheessa yhdenmukaista lähestymistapaa[2]. Tässä lausunnossa käytettävä terminologia on esitetty jäljempänä jaksoissa 2.2 ja 2.3.

4. Irlannin valvontaviranomainen esitti seuraavat kysymykset:

Kysymys 1: Katsotaanko lopullisen tekoälymallin, joka on koulutettu henkilötietoja käyttäen, kaikissa tapauksissa olevan henkilötietojen määritelmän vastainen (yleisen tietosuoja-asetuksen 4 artiklan 1 kohdan mukaisesti)?

Jos vastaus ensimmäiseen kysymykseen on myöntävä:

i. Missä tekoälymalliin johtavien käsittelytoimien vaiheessa henkilötietoja ei enää käsitellä?
   a) Miten voidaan osoittaa, että tekoälymalli ei käsittele henkilötietoja?

ii. Liittyykö asiaan tekijöitä, joiden vuoksi lopullisen tekoälymallin toimintaa ei enää pidetä anonyyminä?
   a) Jos vastaus on kyllä, miten voidaan osoittaa toimenpiteet, joita on toteutettu näiden tekijöiden lieventämiseksi tai ehkäisemiseksi tai niiltä suojautumiseksi (sen varmistamiseksi, että tekoälymallissa ei käsitellä henkilötietoja)?

Jos vastaus ensimmäiseen kysymykseen on kielteinen:

i. Millaisissa olosuhteissa tätä voi tapahtua?
   a) Jos näin tapahtuu, miten voidaan osoittaa toimenpiteet, joilla on varmistettu, että tekoälymallissa ei käsitellä henkilötietoja?

Kysymys 2: Jos rekisterinpitäjä käyttää oikeutettuja etuja henkilötietojen käsittelyn oikeusperusteena tekoälymallin luomiseksi, päivittämiseksi ja/tai kehittämiseksi, miten rekisterinpitäjän olisi osoitettava oikeutettujen etujen asianmukaisuus oikeusperusteena kolmannen osapuolen ja ensimmäisen osapuolen tietojen käsittelyssä?

i. Mitä näkökohtia kyseisen rekisterinpitäjän olisi otettava huomioon sen varmistamiseksi, että niiden rekisteröityjen edut, joiden henkilötietoja käsitellään, ovat asianmukaisesti tasapainossa suhteessa kyseisen rekisterinpitäjän etuihin seuraavien seikkojen osalta:
   a) Kolmannen osapuolen tiedot

   b) Ensimmäisen osapuolen tiedot

Kysymys 3: Jos rekisterinpitäjä käyttää oikeutettuja etuja oikeusperusteena henkilötietojen käsittelylle tekoälymallissa tai tekoälyjärjestelmässä, johon tekoälymalli kuuluu, miten rekisterinpitäjän olisi osoitettava koulutuksen jälkeen oikeutettujen etujen asianmukaisuus oikeusperusteena?

[2] Pyyntö, s. 1.

Kysymys 4: Jos on todettu, että tekoälymalli on luotu, päivitetty tai kehitetty käyttäen laittomasti käsiteltyjä henkilötietoja, mikä on tämän mahdollinen vaikutus tekoälymallin jatkuvan tai myöhemmän käsittelyn tai toiminnan laillisuuteen joko sellaisenaan tai osana tekoälyjärjestelmää, kun

i. tekoälymalli käsitteleekö henkilötietoja joko yksin tai osana tekoälyjärjestelmää tai kun
ii. tekoälymalli tai tekoälyjärjestelmään kuuluva tekoälymalli ei käsittele henkilötietoja?

1.2 Yleisen tietosuoja-asetuksen 64 artiklan 2 kohdan mukaista lausuntoa koskevan pyynnön hyväksyttävyys

5. Yleisen tietosuoja-asetuksen 64 artiklan 2 kohdassa säädetään erityisesti, että jokainen valvontaviranomainen voi pyytää minkä tahansa yleisluonteisen tai useammassa kuin yhdessä jäsenvaltiossa vaikutuksia tuottavan asian käsittelyä tietosuojaneuvostossa lausunnon saamiseksi.

6. Pyynnön esittänyt valvontaviranomainen esitti tietosuojaneuvostolle kysymyksiä tietosuojanäkökohdista tekoälymallien yhteydessä. Se täsmensi pyynnössä, että vaikka monet organisaatiot käyttävät nykyään tekoälymalleja, mukaan lukien laajat kielimallit, niiden toiminta, koulutus ja käyttö herättävät *useita laaja-alaisia tietosuojaan liittyviä huolenaiheita*[3], jotka *vaikuttavat rekisteröityihin kaikkialla EU:ssa/ETA:ssa*[4].

7. Pyynnössä esitetään pääasiassa kysymyksiä i) henkilötietojen käsitteen soveltamisesta; ii) lainmukaisuuden periaatteesta, jossa otetaan erityisesti huomioon oikeutetun edun oikeusperuste, tekoälymallien yhteydessä sekä iii) henkilötietojen laittoman käsittelyn seurauksista tekoälymallien kehittämisvaiheessa, mallin myöhemmästä käsittelystä tai toiminnasta.

8. Näin ollen tämä pyyntö koskee yleisen tietosuoja-asetuksen 64 artiklan 2 kohdassa tarkoitettua *yleisluonteisen asian käsittelyä*. Asia liittyy erityisesti yleisen tietosuoja-asetuksen 4 artiklan 1 kohdan, 5 artiklan 1 kohdan a alakohdan ja 6 artiklan tulkintaan ja soveltamiseen suhteessa henkilötietojen käsittelyyn tekoälymallien kehittämisessä ja käytössä. Kuten pyynnön esittänyt valvontaviranomainen korosti, näiden säännösten soveltaminen tekoälymalleihin herättää systeemisiä, abstrakteja ja uudenlaisia kysymyksiä [5]. Tekoälymallien nopea kehittäminen ja käyttöönotto yhä useammassa organisaatiossa herättää erityisiä kysymyksiä, ja, kuten pyynnössä todetaan, *Euroopan tietosuojaneuvosto hyötyy suuresti yhteisen kannan muodostamisesta tässä pyynnössä esiin tuotuihin kysymyksiin, sillä nämä kysymykset ovat keskeisiä Euroopan tietosuojaneuvoston suunnitellun työn kannalta lyhyellä ja keskipitkällä aikavälillä*[6]. Lisäksi tekoälyteknologiat luovat monia mahdollisuuksia ja hyötyjä monilla eri aloilla ja yhteiskunnallisissa toiminnoissa. Lisäksi yleinen tietosuoja-asetus on oikeudellinen kehys, joka kannustaa vastuulliseen innovointiin. Tästä seuraa, että on yleisen edun mukaista tehdä tämä arviointi tietosuojaneuvoston lausunnon muodossa, jotta voidaan varmistaa tiettyjen yleisen tietosuoja-asetuksen säännösten yhdenmukainen soveltaminen tekoälymallien yhteydessä.

[3] Pyyntö, s. 1.
[4] Ks. edellinen alaviite.
[5] Pyyntö, s. 2.
[6] Pyyntö, s. 1. Kuten Euroopan tietosuojaneuvoston työohjelmassa 2024–2025, joka hyväksyttiin 8. lokakuuta 2024 ja joka on saatavilla osoitteessa https://www.edpb.europa.eu/system/files/2024-10/edpb_work_programme_2024-2025_en.pdf, todetaan, Euroopan tietosuojaneuvosto aikoo antaa *muun muassa* ohjeet anonymisoinnista, pseudonymisoinnista ja tiedonharavoinnista generatiivisen tekoälyn yhteydessä.

9. Yleisen tietosuoja-asetuksen 64 artiklan 2 kohdan mukaisessa vaihtoehtoisessa edellytyksessä viitataan asioihin, jotka *tuottavat vaikutuksia useammassa kuin yhdessä jäsenvaltiossa*. Euroopan tietosuojaneuvosto muistuttaa, että termiä "vaikutukset" on tulkittava *lato sensu*, laajassa merkityksessä, eikä sitä siis rajoiteta pelkästään oikeudellisiin vaikutuksiin[7]. Koska yhä useampia tekoälymalleja koulutetaan ja käytetään yhä useammissa organisaatioissa Euroopan talousalueella, ne vaikuttavat suureen määrään rekisteröityjä kaikkialla Euroopan talousalueella, joista jotkut ovat jo tuoneet esiin huolensa toimivaltaiselle paikalliselle valvontaviranomaiselleen[8]. Näin ollen Euroopan tietosuojaneuvosto katsoo, että myös pyynnön esittäneen valvontaviranomaisen esille ottama asia täyttää tämän ehdon.

10. Pyyntö sisältää kirjalliset perustelut taustoista ja motiiveista, joiden perusteella kysymykset esitetään tietosuojaneuvostolle, mukaan lukien asiaa koskeva oikeudellinen kehys. Sen vuoksi tietosuojaneuvosto katsoo, että pyyntö on perusteltu Euroopan tietosuojaneuvoston työjärjestyksen 10 artiklan 3 kohdan mukaisesti.

11. Yleisen tietosuoja-asetuksen 64 artiklan 3 kohdan mukaisesti tietosuojaneuvosto ei anna lausuntoa, jos se on jo antanut asiasta lausunnon[9]. Euroopan tietosuojaneuvosto ei ole antanut lausuntoa samasta asiasta eikä vielä vastannut pyynnössä esitettyihin kysymyksiin.

12. Näistä syistä tietosuojaneuvosto katsoo, että pyyntö voidaan ottaa käsiteltäväksi ja siitä johtuvia kysymyksiä olisi analysoitava tässä yleisen tietosuoja-asetuksen 64 artiklan 2 kohdan nojalla annetussa lausunnossa, jäljempänä ’lausunto’.

2 Soveltamisala ja keskeiset käsitteet

2.1 Lausunnon soveltamisala

13. Tietosuojaneuvosto on pyynnön esittäneen valvontaviranomaisen kanssa samaa mieltä siitä, että tietosuojan näkökulmasta tekoälymallien kehittäminen ja käyttö herättävät perustavanlaatuisia kysymyksiä tietosuojasta. Kysymykset liittyvät erityisesti seuraaviin asioihin: i) milloin ja miten tekoälymallia voidaan pitää ”anonyyminä” (pyynnön kysymys 1); ii) miten rekisterinpitäjät voivat osoittaa, että oikeutettu etu on asianmukainen oikeusperuste kehittämis- (pyynnön kysymys 2) ja käyttövaiheissa (pyynnön kysymys 3); ja iii) onko henkilötietojen laittomalla käsittelyllä kehittämisvaiheessa vaikutuksia tekoälymallin myöhemmän käsittelyn tai toiminnan laillisuuteen (pyynnön kysymys 4).

14. Tietosuojaneuvosto muistuttaa, että valvontaviranomaiset ovat vastuussa yleisen tietosuoja-asetuksen soveltamisen valvonnasta ja niiden olisi edistettävä sen yhtenäistä soveltamista koko unionissa[10]. Sen vuoksi valvontaviranomaisten toimivaltaan kuuluu tutkia tiettyjä tekoälymalleja ja tehdä tässä yhteydessä tapauskohtaisia arviointeja.

15. Tämä lausunto tarjoaa toimivaltaisille valvontaviranomaisille puitteet sellaisten erityistapausten arvioimiseksi, joissa (jotkin) pyynnössä esitetyt kysymykset nousevat esiin. Tämän lausunnon

[7] Euroopan tietosuojaneuvoston sisäinen asiakirja 3/2019 yleisen tietosuoja-asetuksen 64 artiklan 2 kohtaa koskevista sisäisistä ohjeista, annettu 8. lokakuuta 2019, kohta 15, saatavilla osoitteessa https://www.edpb.europa.eu/system/files/2022-07/internaledpb_document_201903_art64.2_en.pdf.
[8] Pyyntö, s. 1–2.
Yleisen tietosuoja-asetuksen[9] 64 artiklan 3 kohta ja Euroopan tietosuojaneuvoston työjärjestyksen 10 artiklan 4 kohta.
[10] Yleisen tietosuoja-asetuksen 51 artiklan 1 kohta ja 51 artiklan 2 kohta.

tarkoituksena ei ole olla tyhjentävä, vaan pikemminkin esittää yleisiä näkökohtia asianomaisten säännösten tulkinnasta, jotka toimivaltaisten valvontaviranomaisten olisi otettava mahdollisimman tarkasti huomioon tutkintavaltuuksia käyttäessään. Vaikka tämä lausunto on osoitettu toimivaltaisille valvontaviranomaisille ja liittyy niiden toimintaan ja toimivaltaan, se ei vaikuta yleisen tietosuoja-asetuksen mukaisiin rekisterinpitäjien ja henkilötietojen käsittelijöiden velvollisuuksiin. Yleisen tietosuoja-asetuksen 5 artiklan 2 kohdassa vahvistetun osoitusvelvollisuusperiaatteen mukaisesti rekisterinpitäjät vastaavat kaikista henkilötietojen käsittelyä koskevista periaatteista ja osoittavat niiden noudattamisen.

16. Joissakin tapauksissa lausunnossa voidaan antaa joitakin esimerkkejä, mutta ottaen huomioon pyynnössä esitettyjen kysymysten laaja soveltamisala ja siinä käsitellyt erityyppiset tekoälymallit, kaikkia mahdollisia skenaarioita ei käsitellä tässä lausunnossa. Tekoälymalleihin liittyvät teknologiat kehittyvät nopeasti. Niinpä tässä lausunnossa esitettyjä Euroopan tietosuojaneuvoston huomioita olisi tulkittava tämän valossa.

17. **Tässä lausunnossa ei analysoida seuraavia säännöksiä, joilla voi edelleen olla tärkeä rooli arvioitaessa tekoälymalleihin sovellettavia tietosuojavaatimuksia:**

- **Erityisiä henkilötietoryhmiä koskeva käsittely** Tietosuojaneuvosto muistuttaa yleisen tietosuoja-asetuksen 9 artiklan 1 kohdassa säädetystä kiellosta käsitellä erityisiä tietoryhmiä ja rajoitetuista poikkeuksista yleisen tietosuoja-asetuksen 9 artiklan 2 kohdassa [11]. Tältä osin Euroopan unionin tuomioistuin (**CJEU**) selvensi edelleen, että " *tilanteessa, jossa sekä arkaluonteisia tietoja että muita kuin arkaluonteisia tietoja [...] kerätään kokonaisuutena eikä tietoja voida niitä kerättäessä erotella, on katsottava, että tämän tietojoukon käsittely on yleisen tietosuoja-asetuksen 9 artiklan 1 kohdan mukaisesti kiellettyä silloin, kun siihen sisältyy vähintäänkin yksi arkaluonteinen tieto eikä mitään yleisen tietosuoja-asetuksen 9 artiklan 2 kohdassa olevaa poikkeusta voida soveltaa*"[12]. Unionin tuomioistuin korosti myös, että *" yleisen tietosuoja-asetuksen 9 artiklan 2 kohdassa säädetyn poikkeuksen soveltamiseksi on selvitettävä, onko rekisteröity tarkoittanut nimenomaisesti ja toteuttamalla selkeästi suostumusta ilmaisevan toimen saattaa kyseessä olevat henkilötiedot suuren yleisön saataville[13]"*. Nämä seikat olisi otettava huomioon, kun henkilötietojen käsittely tekoälymallien yhteydessä koskee erityisiä tietoryhmiä.

- **Automatisoitu päätöksenteko, mukaan lukien profilointi**: Tekoälymallien yhteydessä suoritettavat käsittelytoimet voivat kuulua yleisen tietosuoja-asetuksen 22 artiklan soveltamisalaan. Kyseisessä artiklassa asetetaan rekisterinpitäjille lisävelvoitteita ja annetaan rekisteröidyille lisäsuojatoimia. Tietosuojaneuvosto muistuttaa tältä osin automaattista

[11] Ks. myös Euroopan tietosuojaneuvoston raportti ChatGPT-työryhmän tekemästä työstä, hyväksytty 23. toukokuuta 2024, 18 kohta: Erityisiin henkilötietoryhmiin kuuluvien tietojen käsittelyyn on lisäksi sovellettava jotakin 9 artiklan 2 kohdassa säädetyistä poikkeuksista, jotta käsittely olisi laillista. *Yksi näistä poikkeuksista voi periaatteessa sisältyä tietosuoja-asetuksen 9 artiklan 2 kohdan e alakohtaan. Pelkästään se, että henkilötiedot ovat julkisesti saatavilla, ei kuitenkaan tarkoita, että rekisteröity on nimenomaisesti saattanut kyseiset tiedot julkisiksi [...]*.

[12] Unionin tuomioistuimen tuomio, 4 päivänä heinäkuuta 2023, asiassa C-252/21, *Meta vastaan Bundeskartellamt* (ECLI:EU:C:2023:537), 89 kohta.

[13] Unionin tuomioistuimen tuomio, 4 päivänä heinäkuuta 2023, asiassa C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), 77 kohta.

yksilöllistä päätöksentekoa ja profilointia koskevista ohjeistaan asetuksen (EU) 2016/679 [14] soveltamiseksi.

- **Tarkoitusten yhteensopivuus**: Yleisen tietosuoja-asetuksen 6 artiklan 4 kohdassa säädetään tiettyjen oikeusperusteiden osalta kriteereistä, jotka rekisterinpitäjän on otettava huomioon varmistaakseen, että käsittely toista tarkoitusta varten sopii yhteen sen tarkoituksen kanssa, jota varten henkilötiedot on alun perin kerätty. Tällä säännöksellä voi olla merkitystä tekoälymallien kehittämisen ja käyttöönoton kannalta, ja valvontaviranomaisten olisi arvioitava sen sovellettavuutta.

- **Tietosuojaa koskeva vaikutustenarviointi** (**DPIA**) (yleisen tietosuoja-asetuksen 35 artikla): Tietosuojaa koskeva vaikutustenarviointi on tärkeä osa osoitusvelvollisuutta, kun tekoälymallien yhteydessä tapahtuva käsittely todennäköisesti aiheuttaa suuren riskin luonnollisten henkilöiden oikeuksille ja vapauksille[15].

- **Sisäänrakennetun tietosuojan periaate** (yleisen tietosuoja-asetuksen 25 artiklan 1 kohta): Sisäänrakennettu tietosuoja on olennainen suojatoimi, jota valvontaviranomaisten tulee arvioida tekoälymallin kehittämisen ja käyttöönoton yhteydessä.

2.2 Keskeiset käsitteet

18. Euroopan tietosuojaneuvosto haluaa ensin selventää terminologiaa ja käsitteitä, joita se käyttää tässä lausunnossa ja vain tämän lausunnon tarkoituksia varten:

- ”**Ensimmäisen osapuolen tiedoilla**” tarkoitetaan henkilötietoja, jotka rekisterinpitäjä on kerännyt rekisteröidyiltä.

- ”**Kolmannen osapuolen tiedoilla**” tarkoitetaan henkilötietoja, joita rekisterinpitäjät eivät ole keränneet rekisteröidyiltä, vaan ne on vastaanotettu kolmannelta osapuolelta, kuten tietojen välittäjältä, tai kerätty käyttämällä verkkosivujen haravointia.

- "**Verkkosivujen haravointi**" on yleisesti käytetty tekniikka tietojen keräämiseksi julkisesti saatavilla olevista verkkolähteistä. Esimerkiksi sanomalehdistä, sosiaalisesta mediasta, keskusteluryhmistä ja henkilökohtaisilta verkkosivustoilta poistetuissa tiedoissa voi olla henkilötietoja.

- Pyynnössä viitataan **tekoälymallien ”elinkaareen”** sekä eri vaiheisiin, jotka liittyvät muun muassa tekoälymallien luomiseen, kehittämiseen, kouluttamiseen, päivittämiseen, hienosäätöön, toimintaan tai myöhempään koulutukseen. Tietosuojavaltuutettu tunnistaa, että olosuhteiden mukaan tekoälymallien kehittämiseen ja käyttöönottoon voi sisältyä tällaisia vaiheita, ja niissä voidaan käsitellä henkilötietoja eri käsittelytarkoituksiin. Tässä lausunnossa Euroopan tietosuojaneuvosto pitää kuitenkin tärkeänä, että todennäköisesti esiintyvien vaiheiden luokittelua selkeytetään. Siten Euroopan tietosuojaneuvosto viittaa tässä lausunnossa ”**kehittämisvaiheeseen**” ja ”**käyttöönottovaiheeseen**”. Tekoälymallin kehittäminen kattaa kaikki

[14] Tietosuojatyöryhmän (**WP29**) Suuntaviivat automatisoiduista yksittäispäätöksistä ja profiloinnista asetuksen (EU) 2016/679 täytäntöön panemiseksi, sellaisena kuin se on viimeksi tarkistettuna ja hyväksyttynä 6. helmikuuta 2018, jonka Euroopan tietosuojaneuvosto vahvisti 25. toukokuuta 2018. Ks. myös unionin tuomioistuimen tuomio C-634/21, 7 päivänä joulukuuta 2023, asiassa C-634/21, *SCHUFA Holding ym.* (ECLI:EU:C:2023:957).

[15] Tietosuojatyöryhmän ohjeet tietosuojaa koskevasta vaikutustenarvioinnista ja keinoista selvittää ”liittyykö käsittelyyn todennäköisesti” asetuksessa (EU) 2016/679 tarkoitettu "korkea riski”, tarkistettu ja hyväksytty 4. lokakuuta 2017.

vaiheet ennen tekoälymallin käyttöönottoa, ja niitä ovat muun muassa koodin kehittäminen, henkilötietojen koulutustarkoiteuksessa, koulutustarkoiteukseen kerättyjen henkilötietojen esikäsittely ja kouluttaminen. Tekoälymallin käyttöönotto kattaa kaikki tekoälymallin käytön vaiheet, ja se voi sisältää myös kaikenlaisia kehittämisvaiheen jälkeen toteutettavia toimia. Tietosuojaneuvosto on edelleen perillä siitä, että käyttötapauksia on monenlaisia ja että niillä voi olla erilaisia seurauksia henkilötietojen käsittelyyn; näin ollen valvontaviranomaisten olisi harkittava, ovatko tässä lausunnossa esitetyt huomautukset merkityksellisiä arvioidun käsittelyn kannalta.

- Tietosuojaneuvosto korostaa myös, että tarvittaessa termi "**koulutus**" viittaa siihen kehittämisvaiheen osaan, jossa tekoälymallit oppivat datasta suorittamaan aiotun tehtävänsä (kuten tämän lausunnon seuraavassa jaksossa selitetään).

- Seuraavassa jaksossa tarkennetaan **tekoälymallien** käsitettä ja soveltamisalaa siten kuin Euroopan tietosuojaneuvosto on ne ymmärtänyt tämän lausunnon tarkoituksiin.

2.3 Tekoälymallit lausunnon yhteydessä

19. EU:n tekoälysäädöksessä (**AI Act**)[16] tekoälyjärjestelmä määritellään seuraavasti: tekoälyjärjestelmällä tarkoitetaan "*konepohjaista järjestelmää, joka on suunniteltu toimimaan käyttöönoton jälkeen vaihtelevilla autonomian tasoilla ja jossa voi ilmetä mukautuvuutta käyttöönoton jälkeen ja joka päättelee vastaanottamastaan syötteestä eksplisiittisiä tai implisiittisiä tavoitteita varten, miten tuottaa tuotoksia, kuten ennusteita, sisältöä, suosituksia tai päätöksiä, jotka voivat vaikuttaa fyysisiin tai virtuaalisiin ympäristöihin*" [17] . Tekoälysäädöksen johdanto-osan 12 kappaleessa selitetään tarkemmin "tekoälyjärjestelmän" käsitettä. Siten yksi tekoälyjärjestelmien keskeisistä piirteistä on niiden päättelykyky. Tekoälyjärjestelmää kehitettäessä käytettäviä päättelyn mahdollistavia tekniikoita ovat esimerkiksi koneoppimismenetelmät ja logiikkaan ja tietämykseen perustuvat menetelmät.

20. Toisaalta tekoälymallit määritellään tekoälysäädöksessä vain epäsuorasti: ”*Vaikka tekoälymallit ovat tekoälyjärjestelmien olennaisia osia, ne eivät yksinään muodosta tekoälyjärjestelmiä. Tekoälyjärjestelmät edellyttävät lisäkomponenttien, kuten käyttöliittymän, lisäämistä tekoälymalleihin. Tekoälymallit tyypillisesti integroidaan tekoälyjärjestelmiin ja ovat osa niitä*”[18].

21. Euroopan tietosuojaneuvosto ymmärtää, että pyynnössä ehdotettu tekoälymallin määritelmä on suppeampi kuin tekoälysäädöksessä esitetty määritelmä, koska pyynnössä tekoälymalliin viitataan siten, että sillä *tarkoitetaan tuotetta, joka on saatu koulutusdataan sovellettavista koulutusmekanismeista tekoälyn, koneoppimisen, syväoppimisen tai muiden asiaan liittyvien käsittely-ympäristöjen yhteydessä*, ja täsmennetään lisäksi, että käsitettä *sovelletaan tekoälymalleihin, jotka on tarkoitettu lisäkoulutukseen, hienosäätöön ja/tai kehittämiseen, sekä tekoälymalleihin, jotka eivät ole*.[19]

22. Tämän pohjalta Euroopan tietosuojaneuvosto hyväksyi tämän lausunnon edellyttäen, että tekoälyjärjestelmä tukeutuu tekoälymalliin halutun tavoitteen saavuttamiseksi sisällyttämällä mallin

[16] Euroopan parlamentin ja neuvoston asetus (EU) 2024/1689, annettu 13 päivänä kesäkuuta 2024, tekoälyä koskevista yhdenmukaistetuista säännöistä ja asetusten (EY) N:o 300/2008, (EU) N:o 167/2013, (EU) N:o 168/2013, (EU) 2018/858, (EU) 2018/1139 ja (EU) 2019/2144 sekä direktiivien 2014/90/EU, (EU) 2016/797 ja (EU) 2020/1828 muuttamisesta (tekoälysäädös).
[17] Tekoälysäädöksen 3 artiklan 1 kohta.
[18] Tekoälysäädöksen johdanto-osan 97 kappale.
[19] Pyyntö, s. 3.

laajempaan kehykseen (esim. asiakaspalvelua koskeva tekoälyjärjestelmä voisi käyttää tekoälymallia, joka on koulutettu aiempien keskustelutietojen perusteella, jotta se voisi vastata käyttäjien kyselyihin).

23. Lisäksi tämän lausunnon kannalta merkityksellisiä ovat tekoälymallit (tai "**mallit**"), jotka on kehitetty koulutusprosessissa. Tällainen koulutusprosessi on osa kehittämisvaihetta, jossa mallit oppivat tiedoista aiotun tehtävän suorittamiseksi. Koulutusprosessi edellyttää siis tietokokonaisuutta, josta tekoälymalli tunnistaa ja "oppii" toimintamalleja. Näissä tapauksissa malli käyttää erilaisia tekniikoita rakentaakseen koulutustietoaineistosta poimitusta tiedosta esityksen. Tämä koskee erityisesti koneoppimista.

24. Käytännössä mikä tahansa tekoälymalli on algoritmi, jonka toiminta määräytyy tiettyjen elementtien perusteella. Esimerkiksi syväoppimismallit ovat usein neuroverkkoja, joissa on useita kerroksia. Ne koostuvat keinotekoisista toisiinsa kytkeytyneistä neuroneista. Niillä on kytkentäkohtainen paino, joita voidaan muuttaa koulutuksessa syötteiden ja tuotosten välisten suhteiden opettamiseksi. Yksinkertaisen syväoppimismallin tunnusmerkkejä ovat i) kunkin kerroksen tyyppi ja koko, ii) kullekin reunalle annettu paino (kutsutaan joskus parametreiksi), iii) aktivointifunktiot[20] kerrosten välillä ja iv) mahdollisesti muut operaatiot, joita voi tapahtua kerrosten välillä. Kun esimerkiksi koulutetaan yksinkertaista syväoppimismallia kuvan luokittelua varten, syötteet (**kuvapikselit**) yhdistetään tuotoksiin ja painoja voidaan mukauttaa siten, että suurimman osan ajasta saadaan oikea tuotos.

25. Muita esimerkkejä syväoppimismalleista ovat laajat kielimallit ja generatiivinen tekoäly, joita käytetään esimerkiksi ihmisen tekemää muistuttavan sisällön tuottamiseen ja uuden datan luomiseen.

26. **Edellä esitettyjen näkökohtien perusteella ja pyynnön mukaisesti tämän lausunnon soveltamisala kattaa vain sellaiset tekoälymallit, jotka on koulutettu henkilötietojen avulla.**

3 Pyynnön perusteltavuus

3.1 Tekoälymallien luonne ja henkilötietojen määritelmä

27. Yleisen tietosuoja-asetuksen 4 artiklan 1 kohdan määritelmän mukaan henkilötiedoilla tarkoitetaan ”*kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, (jäljempänä ’rekisteröity’) liittyviä tietoja*”. Lisäksi yleisen tietosuoja-asetuksen johdanto-osan 26 kappaleessa säädetään, että tietosuojaperiaatteita ei pitäisi soveltaa anonyymeihin tietoihin eli tietoihin, jotka eivät liity tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, kun otetaan huomioon ”*kaikki keinot, joita rekisterinpitäjä tai muu henkilö voi kohtuullisen todennäköisesti käyttää*”. Tähän sisältyy seuraavaa: i) tiedot, jotka eivät ole koskaan liittyneet tunnistettuun tai tunnistettavissa olevaan henkilöön, ja ii) henkilötiedot, jotka on muutettu anonyymeiksi siten, että rekisteröity ei ole tai ei ole enää tunnistettavissa.

28. Näin ollen pyynnön kysymykseen i)[21] voidaan vastata analysoimalla, olisiko henkilötietojen käsittelyä sisältävän koulutuksen tuloksena syntyvää tekoälymallia pidettävä kaikissa tapauksissa anonyyminä. Kysymyksen sanamuodon perusteella tietosuojaneuvosto viittaa tässä jaksossa tekoälymallin ”koulutusprosessiin”.

[20] Toisin sanoen toiminnot laskevat syötteiden ja painojen perusteella keinotekoisen neuronin tuotoksen, joka lähetetään neuroverkon seuraavaan kerrokseen.

[21] *Katsotaanko, että lopullisessa tekoälymallissa, joka on koulutettu henkilötietojen avulla, henkilötietojen määritelmä ei täyty kaikissa tapauksissa (GDPR:n 4 artiklan 1 kohdan mukaisesti)?*

29. Euroopan tietosuojaneuvosto haluaa ennen kaikkea esittää seuraavat yleiset huomiot. Tekoälymallit on yleensä suunniteltu tekemään ennusteita tai johtopäätöksiä siihen katsomatta, onko niitä koulutettu henkilötietoja käsittävällä data-aineistolla vai ei. Toisin sanoen ne on suunniteltu tekemään päätelmiä. Lisäksi tekoälymallit, joita on koulutettu henkilötiedoilla, on usein suunniteltu tekemään päätelmiä henkilöistä, jotka ovat erilaisia kuin ne, joiden henkilötietoja on käytetty tekoälymallin kouluttamisessa. Jotkin tekoälymallit on kuitenkin nimenomaan suunniteltu antamaan henkilötietoja niistä henkilöistä, joiden henkilötietoja on käytetty mallin kouluttamisessa, tai asettamaan jollakin tavalla tällaiset tiedot saataville. Tällaiset tekoälymallit sisältävät luontaisesti (ja tyypillisesti väistämättä) tietoja, jotka liittyvät tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, joten niissä on kyse henkilötietojen käsittelystä. Siksi tämäntyyppisiä tekoälymalleja ei voida pitää anonyymeina. Tämä koskee esimerkiksi i) generatiivista mallia, joka on hienosäädetty äänitallenteiden perusteella jäljittelemään henkilön ääntä, tai ii) mallia, joka on suunniteltu vastaamaan koulutuksessa annettuja henkilötietoja hyödyntäen, kun pyydetään tiettyä henkilöä koskevia tietoja.

30. Vastatessaan pyynnön kysymykseen 1 edellä mainittujen näkökohtien perusteella Euroopan tietosuojaneuvosto keskittyy sellaisiin tekoälymalleihin, joita ei ole suunniteltu antamaan koulutuksen yhteydessä käytettyjä henkilötietoja.

31. Tietosuojaneuvosto katsoo, että vaikka tekoälymallia ei ole tarkoituksellisesti suunniteltu tuottamaan tietoja, jotka liittyvät käytetyn koulutusaineiston tunnistettuun tai tunnistettavaan luonnolliseen henkilöön, koulutusaineiston tiedot, kuten henkilötiedot, voivat silti absorboitua mallin parametreihin, ja ne voidaan esittää matemaattisten objektien avulla. Ne voivat poiketa alkuperäisistä koulutuksen datapisteistä, mutta niissä voi silti säilyä näiden tietojen alkuperäinen informaatio, joka voidaan poimia tai saada muulla tavoin suoraan tai epäsuoraan mallista. Jos tunnistetuista tai tunnistettavissa olevista henkilöistä, joiden henkilötietoja on käytetty mallin kouluttamiseen, voidaan saada tietoja tekoälymallista keinoin, jotka ovat kohtuullisen todennäköisesti käytettävissä, voidaan päätellä, että tällainen malli ei ole anonyymi.

32. Tältä osin pyynnössä todetaan, että *olemassa olevissa tutkimusjulkaisuissa tuodaan esiin joitakin mahdollisia haavoittuvuuksia, joita tekoälymalleissa voi olla ja jotka voivat johtaa henkilötietojen käsittelyyn,*[22] *sekä henkilötietojen käsittelyyn, jota voi tapahtua, kun malleja otetaan käyttöön käytettäväksi muiden tietojen kanssa joko ohjelmointirajapintojen (**API**) tai nopeiden rajapintojen*[23] kautta.

33. Samaan tapaan koulutustiedon poimimista koskeva tutkimus on erityisen dynaamista[24]. Se osoittaa, että joissakin tapauksissa on mahdollista käyttää keinoja, joilla voidaan kohtuullisen todennäköisesti

[22] Tällaisia ovat esimerkiksi joukkoon kuulumiseen perustuvaa päättelyä hyödyntävät hyökkäykset (OWASP) ja mallin inversiota hyödyntävät hyökkäykset – Model Inversion Attacks (OWASP & Veale et al, 2018).
[23] Pyyntö, s. 1–2.
[24] Ks. tältä osin esimerkiksi: i) Veale Michael, Binns Reuben ja Edwards Lilian, 2018 *Algorithms that remember: model inversion attacks and data protection law*. Phil. Trans. R. Soc. A 376: 20180083, saatavilla osoitteessa http://dx.doi.org/10.1098/rsta.2018.0083; ii) Brown H., Lee K., Mireshghallah F., Shokri R., and Tramèr F., *What Does it Mean for a Language Model to Preserve Privacy?*, 2022, ACM Digital Library, FAccT '22, June 20, 2022, Seoul, Republic of Korea, osoitteessa https://dl.acm.org/doi/abs/10.1145/3531146.3534642; iii) Vassilev A., Oprea A., Fordyce A., Anderson H., *Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations*, January 2024, National Institute of Standards and Technology, osoitteessa https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2023.pdf; iv) Carlini N., Tramèr F., Wallace E., Jagielski M., Herbert-Voss A., Lee K., Roberts A., Brown T., Song D., Erlingsson U., Oprea A., Raffel C., *Extracting Training Data from Large Language Models*, arXiv:2012.07805v2 [cs.CR] 15.6.2021, osoitteessa https://arxiv.org/pdf/2012.07805; v) Fredrikson M., Jha S., Ristenpart T., *Model Inversion Attacks that Exploit*

poimia henkilötietoja joistakin tekoälymalleista, tai yksinkertaisesti saada henkilötietoja vahingossa vuorovaikutukusessa tekoälymallin kanssa (esimerkiksi tekoälyjärjestelmän osana). Jatkuvat tutkimus alalla auttaa arvioimaan edelleen regurgitaation [25] ja henkilötietojen poimimisen jäännösriskejä kussakin tapauksessa.

34. **Edellä esitettyjen näkökohtien perusteella tietosuojaneuvosto katsoo, että henkilötietojen pohjalta koulutettuja tekoälymalleja ei voida kaikissa tapauksissa pitää anonyymeinä. Sen tekoälymallin määrittämistä anonyymiksi olisi arvioitava tapauskohtaisesti erityisten kriteerien perusteella.**

3.2 Olosuhteet, joissa tekoälymalleja voitaisiin pitää anonyymeina, ja niihin liittyvä demonstrointi

35. Pyynnön [26] kysymyksen 1 osalta Euroopan tietosuojaneuvostoa pyydetään selventämään, missä olosuhteissa henkilötietojen avulla koulutettua tekoälymallia voidaan pitää anonyyminä. Pyynnön [27] kysymyksen 1 kohdan i) alakohdan a alakohdan osalta Euroopan tietosuojaneuvostoa pyydetään selventämään, mitkä todisteet ja/tai asiakirjat valvontaviranomaisten olisi otettava huomioon arvioidessaan, onko tekoälymalli anonyymi.

3.2.1 Anonymisointia koskevat yleiset näkökohdat tarkasteltavana olevassa asiayhteydessä

36. Ilmaisun "*kaikki - - tiedot*" käyttäminen yleisen tietosuoja-asetuksen 4 artiklan 1 kohdan *henkilötietojen* määritelmässä kuvastaa pyrkimystä antaa käsitteelle laaja soveltamisala, joka kattaa kaikki sellaiseen rekisteröityyn "*liittyvät*" tiedot, joka on tunnistettu tai joka voidaan tunnistaa suoraan tai välillisesti.

37. Tieto voi liittyä luonnolliseen henkilöön, vaikka se olisi teknisesti järjestetty tai koodattu (esimerkiksi ainoastaan koneellisesti luettavaan muotoon, olipa se sitten suojattu tai avoin) tavalla, joka ei tee yhteyttä kyseiseen luonnolliseen henkilöön välittömästi ilmeiseksi. Tällaisissa tapauksissa voidaan käyttää ohjelmistosovelluksia, joiden avulla voidaan helposti yksilöidä, tunnistaa ja poimia tiettyjä tietoja. Tämä koskee erityisesti tekoälymalleja, joissa parametrit edustavat tilastollisia suhteita koulutuksessa käytettyjen tietojen välillä ja joissa saattaa olla mahdollista poimia tarkkoja tai epätarkkoja (tilastollinen päättely) henkilötietoja joko suoraan malliin sisältyvien tietojen välisistä suhteista tai tekemällä hakuja kyseiseen malliin.

38. Koska tekoälymalleihin ei yleensä sisälly tietueita, jotka voivat olla suoraan eristettyjä tai linkitettyjä, vaan pikemminkin parametreja, jotka kuvaavat todennäköisiä suhteita mallin sisältämien tietojen välillä, mallin tiedoista voi olla mahdollista tehdä päätelmiä [28] esimerkiksi joukkoon kuulumiseen

---

Confidence Information and Basic Countermeasures, ACM Digital Library, 12 October 2015, osoitteessa https://dl.acm.org/doi/abs/10.1145/2810103.2813677; vi) Zhang Y., Jia R., Pei H., Wang W., Li B., Song D., *The Secret Revealer: Generative Model-Inversion Attacks Against Deep Neural Networks*, arXiv:1911.07135v2 [cs.LG] 18 Apr 2020, osoitteessa https://arxiv.org/pdf/1911.07135.

[25] Generatiiviseen tekoälyyn perustuvassa tekoälyjärjestelmässä regurgitaatio vastaa tilannetta, jossa tuotokset liittyvät suoraan koulutustietoihin.

[26] *Missä olosuhteissa tätä voi tapahtua?*

[27] *Jos näin on, miten osoitetaan toimenpiteet, joilla on varmistettu, että tekoälymalli ei käsittele henkilötietoja?*

[28] (i) Carlini N., Chien S., Nasr M., Song S., Terzis A., Tramer F., *Membership Inference Attacks From First Principles,* arXiv:2112.03570, osoitteessa https://arxiv.org/abs/2112.03570

(ii) Creţu A.M, Guépin F., and De Montjoye Y.A, *Correlation inference attacks against machine learning models*. Health, B19, pp. Adv.10, eadj9260(2024). Doi:10.1126/sciadv.adj9260, osoitteessa https://www.science.org/doi/10.1126/sciadv.adj9260

(iii) Dana L., Pydi M. S., Chevaleyre Y., *memorisation in Attention-only Transformers* arXiv:2411.10115v1 [cs.AI] 15. marraskuuta 2024, osoitteessa: https://arxiv.org/abs/2411.10115

perustuvaa päättelyn perusteella realistisissa skenaarioissa. Jotta valvontaviranomainen voisi sopia rekisterinpitäjän kanssa siitä, että tiettyä tekoälymallia voidaan pitää anonyyminä, sen olisi tarkistettava ainakin, onko se saanut riittävästi näyttöä kohtuullisin keinoin siitä, että (i) koulutustietoihin liittyviä henkilötietoja ei voida poimia[29] mallista; ja ii) kyselyn yhteydessä saadut mallin tuotokset eivät liity rekisteröityihin, joiden henkilötietoja on käytetty mallin kouluttamiseen.

39. Valvontaviranomaisten olisi otettava huomioon kolme seikkaa arvioidessaan, täyttyvätkö nämä edellytykset.

40. Valvontaviranomaisten olisi ensinnäkin otettava huomioon seikat, jotka on yksilöity tietosuojatyöryhmän viimeisimmissä lausunnoissa ja/tai tietosuojaneuvoston asiaa koskevissa ohjeissa. Anonymisoinnin osalta tämän lausunnon antamispäivänä valvontaviranomaisten olisi otettava huomioon tietosuojatyöryhmän lausunnossa 05/2014 anonymisointitekniikoista esitetyt seikat, joissa todetaan, että jos oletettavasti anonyymiksi katsotusta tietoaineistosta ei ole mahdollista erottaa, linkittää ja päätellä tietoja, tietoja voidaan pitää anonyymeinä[30]. Siinä todetaan myös, että "*[J]os ehdotus ei täytä jotakin näistä kriteereistä, tunnistamisriskit olisi arvioitava perusteellisesti*".[31]. **Edellä mainitun poimimisen ja päätelmän todennäköisyyden vuoksi tietosuojaneuvosto katsoo, että tekoälymallit vaativat hyvin todennäköisesti tällaista tunnistusriskien perusteellista arviointia.**

41. Toiseksi tässä arvioinnissa olisi otettava huomioon *kaikki kohtuullisesti toteutettavissa olevat keinot, joita joko rekisterinpitäjä tai joku muu voi kyseisen henkilön tunnistamiseksi käyttää*[32], ja näiden keinojen määrittelyn olisi perustuttava objektiivisiin tekijöihin, kuten tietosuoja-asetuksen johdanto-osan 26 kappaleessa selitetään ja joihin voi kuulua
    a. itse koulutusdatan ominaisuudet, tekoälymalli ja koulutusmenettely[33]

(iv) Gehrke M., Liebenow J., Mohammadi E. & Braun T. et al. *Lifting in Support of Privacy-Preserving Probabilistic Inference*. Künstl Intell, 13. kesäkuuta 2024, osoitteessa https://doi.org/10.1007/s13218-024-00851-y
(v) Hu H., *Membership Inference Attacks and Defenses on Machine Learning Models Literature*, saatavilla osoitteessa: https://github.com/HongshengHu/membership-inference-machine-learning-literature
(vi) Nasr M., Carlini N., Hayase J., Jagielski M., Cooper A. F., Ippolito D., Choquette-Choo C. A., Wallace E., Tramèr F. ja Lee K., *Scalable Extraction of Training Data from (Production) Language Models*, arXiv:2311.17035 28. marraskuuta 2023, osoitteessa https://arxiv.org/abs/2311.17035
(vii) Shokri R., Stronati M., Song C., Shmatikov V., *Membership Inference Attacks against Machine Learning Models* arXiv:1610.05820v2 [cs.CR], 31. maaliskuuta 2017, osoitteessa https://arxiv.org/abs/1610.05820
(viii) Staab R., Vero M., Mislav Balunović, Martin Vechev, 2024, *Beyond Memorization: Violating Privacy Via Inference with Large Language Models*, arXiv:2310.07298v2, 6. toukokuuta 2024, osoitteessa https://arxiv.org/abs/2310.07298
(ix) Wu F., Cui L., Yao S., Yu S., *InferenceAttacks in Machine Learning as a Service: A Taxonomy, Review and Promising Directions arXiv*:2406.02027v1 [cs.LG], 27. kesäkuuta 2024, osoitteessa https://arxiv.org/abs/2406.02027v1
(x) Zhang J., Das D., Kamath G., Tramèr F., *Membership Inference Attacks Cannot Prove that a Model Was Trained On Your Data* arXiv:2409.19798v1, [cs.LG], 29. syyskuuta 2024, osoitteessa https://arxiv.org/abs/2409.19798
(xi) Zhou Z., Xiang J., Chen C. ja Su S., *Quantifying and Analyzing Entity-Level Memorization in Large Language Models*, arXiv:2308.15727v2 [cs.CL] 5 Nov 2023, osoitteessa https://arxiv.org/abs/2308.15727.

[29] Tietojen poiminta käsittää erityisesti tapaukset, joissa henkilötietoja johdetaan tekoälymallista ja joissa käyttöliittymiä käytetään vain vähän tai ei lainkaan.
[30] Tietosuojatyöryhmän lausunto 5/2014, sivu 24.
[31] Tietosuojatyöryhmän lausunto 5/2014, sivu 24.
[32] Unionin tuomioistuimen tuomio 19.10.2016 asiassa C-582/14, *Breyer v. Bundesrepublik Deutschland* (ECLI:EU:C:2016:779), 43 kohta.
[33] Tähän sisältyviä ominaisuuksia ovat esimerkiksi tietueiden yksilöllisyys koulutustiedoissa, tietojen tarkkuus, aggregointi, satunnaistaminen ja erityisesti se, miten ne vaikuttavat tunnistamisalttiuteen.

b. konteksti, jossa tekoälymalli julkaistaan ja/tai jossa sitä käsitellään[34]
c. lisätiedot, jotka mahdollistaisivat tunnistamisen ja jotka voivat olla kyseisen henkilön saatavilla
d. kustannukset ja aika, jonka henkilö tarvitsisi hankkiakseen kyseiset lisätiedot (elleivät ne jo ole saatavilla)[35] ja
e. käsittelyajankohtana saatavilla oleva teknologia sekä tekninen kehitys[36].

42. Kolmanneksi valvontaviranomaisten olisi harkittava, ovatko rekisterinpitäjät arvioineet sitä, että rekisterinpitäjään ja erityyppisiin *"muihin henkilöihin"*, mukaan lukien tekoälymallia käyttävät tahattomat kolmannet osapuolet, sisältyy riski tunnistamisen mahdollisuudesta. Lisäksi olisi harkittava, voidaanko näiden toimijoiden kohtuudella katsoa voivan saada pääsy kyseisiin tietoihin tai käsitellä niitä.

43. **Yhteenvetona voidaan todeta tietosuojaneuvoston katsovan, että jotta tekoälymallia voitaisiin pitää anonyyminä käyttäen kohtuullisia keinoja, i) todennäköisyyden siitä, että henkilötietoja kerätään suoraan (myös probabilistisuus) sellaisista henkilöistä, joiden henkilötietoja on käytetty mallin kouluttamiseen, ja ii) todennäköisyyden siitä, että tällaisia henkilötietoja hankitaan tarkoituksellisesti kyselyistä, olisi oltava merkityksetön[37] kaikille rekisteröidyille**. **Valvontaviran omaisten olisi oletusarvoisesti otettava huomioon, että tekoälymallit edellyttävät todennäköisesti tunnistamisen todennäköisyyden perusteellista arviointia, jotta voidaan tehdä päätelmä niiden mahdollisesta anonyymista luonteesta. Tätä todennäköisyyttä arvioitaessa olisi otettava huomioon *kaikki keinot, joita rekisterinpitäjä tai muu henkilö voi kohtuullisen todennäköisesti käyttää,* ja siinä olisi myös otettava huomioon mallin tahaton (uudelleen)käyttö tai julkistaminen.**

3.2.2 Tunnistamisen jäännöstodennäköisyyden arvioinnissa käytettävät elementit

44. Sekä kehitys- että käyttöönottovaiheessa voidaan toteuttaa toimenpiteitä, joilla vähennetään todennäköisyyttä saada henkilötietoja tekoälymallista, tekoälymallin anonymiteetin arvioinnissa olisi kuitenkin otettava huomioon myös suora pääsy malliin.

45. Lisäksi valvontaviranomaisten olisi tapauskohtaisesti arvioitava, ovatko rekisterinpitäjän toteuttamat toimenpiteet, joilla varmistetaan ja osoitetaan, että tekoälymalli on anonyymi, asianmukaisia ja tehokkaita.

46. Valvontaviranomaisen arvioinnin päätelmissä voi olla eroja erityisesti sellaisen julkisesti saatavilla olevan sellaisen tekoälymallin, jonka käyttäjämäärä on tuntemattoman ja jonka menetelmiä henkilötietojen etsimiseen ja poimimiseen ei tunneta, ja sisäisessä käytössä olevan tekoälymallin välillä, joka on vain työntekijöiden käytettävissä. Vaikka valvontaviranomaisten olisi molemmissa tapauksissa tarkistettava, että rekisterinpitäjät ovat täyttäneet 5 artiklan 2 kohdan ja yleisen tietosuoja-asetuksen 24 artiklan mukaisen vastuuvelvollisuutensa, *keinot, joita muut henkilöt todennäköisesti käyttävät*, voivat vaikuttaa tarkasteltavien mahdollisten skenaarioiden laajuuteen ja

[34] Tämä sisältää asiayhteyteen liittyviä tekijöitä, kuten pääsyn rajoittamisen vain joihinkin henkilöihin ja oikeudellisiin suojatoimiin.
[35] Unionin tuomioistuimen tuomio 7. maaliskuuta 2024 asiassa C-479/22 P, *OC v. Euroopan komissio* (ECLI:EU:C:2024:215), 50 kohta.
[36] Unionin tuomioistuimen tuomio 7. maaliskuuta 2024 asiassa C-479/22 P, *OC v. Euroopan komissio* (ECLI:EU:C:2024:215), 50 kohta.
[37] Unionin tuomioistuimen tuomio 19.10.2016 asiassa C-582/14, *Breyer v. Bundesrepublik Deutschland* (ECLI:EU:C:2016:779), 46 kohta ja unionin tuomioistuimen tuomio 7. maaliskuuta 2024 asiassa C-479/22 P, *OC v. Euroopan komissio* (ECLI:EU:C:2024:215), 51 kohta.

luonteeseen. Sen vuoksi mallin kehittämisen ja käyttöönoton kontekstista riippuen valvontaviranomaiset voivat harkita eri tasoisia testejä ja kykyä kestää hyökkäyksiä.

47. Tältä osin tietosuojaneuvosto esittää jäljempänä ohjeellisen ja ei-tyhjentävän luettelon mahdollisista seikoista, jotka valvontaviranomaiset voivat ottaa huomioon arvioidessaan rekisterinpitäjän anonymiteettiä koskevaa vaatimusta. Muut lähestymistavat voivat olla mahdollisia, jos ne tarjoavat vastaavan suojan tason, erityisesti tekniikan taso huomioon ottaen.

48. Jäljempänä lueteltujen tekijöiden olemassaolo tai puuttuminen ei ole ratkaiseva kriteeri tekoälymallin anonymiteetin arvioinnissa.

3.2.2.1 Tekoälymallin suunnittelu

49. Tekoälymallien suunnittelun osalta valvontaviranomaisten olisi arvioitava rekisterinpitäjien kehittämisvaiheessa omaksumia lähestymistapoja. Tässä yhteydessä olisi tarkasteltava neljän keskeisen alan (jotka on yksilöity jäljempänä) soveltamista ja tehokkuutta.

Lähteiden valinta

50. Ensimmäisellä arviointialueella tarkastellaan tekoälymallin kouluttamiseen käytettyjen lähteiden valintaa. Tähän sisältyy valvontaviranomaisten suorittama arviointi kaikista toimista, joita on toteutettu henkilötietojen keräämisen välttämiseksi tai rajoittamiseksi, mukaan lukien muun muassa i) valintaperusteiden asianmukaisuus, ii) valittujen lähteiden merkityksellisyys ja riittävyys aiottuun tarkoitukseen (tarkoituksiin) nähden ja iii) se, onko epäasianmukaiset lähteet suljettu pois.

Tietojen valmistelu ja minimointi

51. Toinen arvioinnin osa-alue liittyy tietojen valmisteluun koulutusvaihetta varten. Valvontaviranomaisten olisi tarkasteltava erityisesti seuraavia seikkoja: i) onko anonyymien tietojen ja/tai pseudonymisoitujen henkilötietojen käyttöä harkittu ja ii) jos on päätetty olla käyttämättä tällaisia toimenpiteitä, perustelut päätökselle ottaen huomioon aiottu tarkoitus; iii) tietojen minimoinnin istrategiat ja tekniikat, joita on käytetty koulutusprosessiin sisältyvien henkilötietojen määrän rajoittamiseksi; ja iv) ennen mallin koulutusta toteutetut tietoja suodattavat prosessit epäolennaisten henkilötietojen poistamiseksi.

Koulutusta koskevat metodologiset valinnat

52. Kolmas arviointialue koskee luotettavien menetelmien valitsemista tekoälymallien kehittämisessä. Valvontaviranomaisten olisi arvioitava menetelmään liittyviä valintoja, joilla voidaan vähentää tai poistaa merkitsevästi tunnistettavuutta, mukaan lukien muun muassa seuraavat: (i) käytetäänkö menetelmässä ragularisointimetodeja mallin yleistämisen parantamiseksi ja ylisovittamisen vähentämiseksi, ja mikä ratkaisevaa, ii) onko rekisterinpitäjän käytössä asianmukaisia ja tehokkaita yksityisyyttä suojaavia tekniikoita (esim. differentiaalinen yksityisyys).

Mallin tuotoksia koskevat toimenpiteet

53. Viimeinen arviointialue koskee kaikkia itse tekoälymalliin lisättyjä menetelmiä tai toimenpiteitä, jotka eivät välttämättä vaikuta riskiin siitä, että kuka tahansa voi poimia henkilötietoja suoraan mallista, mutta jotka saattavat vähentää todennäköisyyttä saada koulutustietoihin liittyviä henkilötietoja kyselyistä.

3.2.2.2 Tekoälymallin analyysi

54. Arvioidakseen suunnitellun tekoälymallin anonymisoinnin luotettavuutta valvontaviranomaiset varmistavat ensin, että malli on kehitetty suunnitellusti ja että sen tekninen hallinta on tehokas. Valvontaviranomaisten olisi arvioitava, ovatko rekisterinpitäjät suorittaneet asiakirjoihin perustuvia tarkastuksia (sisäisiä tai ulkoisia), joihin sisältyy valittujen toimenpiteiden ja niiden vaikutusten

arviointi tunnistamisen todennäköisyyden rajoittamiseksi. Tähän voisi sisältyä koodikatselmuksista laadittujen raporttien analysointi sekä teoreettinen analyysi, jossa dokumentoidaan niiden toimenpiteiden asianmukaisuus, jotka on valittu vähentämään kyseisen mallin uudelleentunnistamisen todennäköisyyttä.

3.2.2.3 Tekoälymallin testaus ja kyky kestää hyökkäyksiä

55. Lisäksi valvontaviranomaisten olisi otettava huomioon rekisterinpitäjän mallilla suorittamien testien laajuus, toistumistiheys, määrä ja laatu. Valvontaviranomaisten olisi erityisesti otettava huomioon, että onnistunut testaus, joka kattaa laajalti tunnetut viimeisintä tekniikkaa edustavat hyökkäykset, voi ainoastaan olla todiste näiden hyökkäysten vastustuskyvystä. Tämän lausunnon antamispäivään mennessä tähän voisi sisältyä muun muassa strukturoitu testaus seuraavien tekijöiden suhteen: (i) attribuutit ja joukkoon kuulumiseen perustuva päättely, (ii) eksfiltraatio, (iii) koulutustietojen regurgitaatio, (iv) mallin invertointi tai (v) hyökkäysten rekonstruktointi.

3.2.2.4 Dokumentointi

56. Yleisen tietosuoja-asetuksen 5, 24, 25 ja 30 artiklassa ja, jos rekisteröityjen oikeuksiin ja vapauksiin kohdistuu todennäköisesti suuri riski, yleisen tietosuoja-asetuksen 35 artiklassa edellytetään, että rekisterinpitäjät dokumentoivat käsittelytoimensa asianmukaisesti. Tätä sovelletaan myös kaikkeen käsittelyyn, jossa tekoälymallia koulutetaan, vaikka käsittelyn tavoitteena olisi anonymisointi. Valvontaviranomaisten olisi otettava huomioon tällaiset asiakirjat ja rekisterinpitäjien suorittamaan käsittelyyn liittyvien riskien säännöllinen arviointi, jotka ovat perustavanlaatuisia toimia ja osoittavat ettei henkilötietoja käsitellä.

57. **Tietosuojaneuvosto katsoo, että valvontaviranomaisten olisi otettava huomioon asiakirjat aina, kun tiettyä tekoälymallia koskeva anonymiteettihakemus on arvioitava. Tietosuojavaltuutettu huomauttaa, että jos valvontaviranomainen ei pysty anonymiteettipyynnön arvioinnin jälkeen, myöskään asiakirjojen perusteella, vahvistamaan, että tekoälymallin anonymisoimiseksi on toteutettu tehokkaita toimenpiteitä, valvontaviranomainen voi katsoa, että rekisterinpitäjä ei ole täyttänyt yleisen tietosuoja-asetuksen 5 artiklan 2 kohdan mukaista osoitusvelvollisuuttaan. Siten huomioon pitäisi ottaa myös yleisen tietosuoja-asetuksen muiden säännösten noudattaminen.**

58. Ihannetapauksessa valvontaviranomaisten olisi tarkistettava, sisältävätkö rekisterinpitäjän asiakirjat

a. kaikki tiedot, jotka liittyvät tietosuojan vaikutustenarviointeihin, mukaan lukien mahdolliset arvioinnit ja päätökset, joiden perusteella määritellään, ettei tietosuojan vaikutustenarviointi ole ollut tarpeen;

b. kaiken tietosuojavastaavan (jos tietosuojavastaava on nimitetty tai olisi pitänyt nimittää) antaman neuvonnan tai palautteen;

c. tiedot tekoälymallin suunnittelun yhteydessä toteutetuista teknisistä ja organisatorisista toimenpiteistä, joilla vähennetään tunnistamisen todennäköisyyttä, mukaan lukien uhkamalli ja riskinarvioinnit, joihin nämä toimenpiteet perustuvat. Tähän tulisi sisältyä kutakin koulutusaineiston lähdettä koskevat erityistoimenpiteet, mukaan lukien asiaankuuluvat lähteiden URL-osoitteet ja kuvaukset toteutetuista (tai kolmannen osapuolen tietokokonaisuuksien tarjoajien jo toteuttamista) toimenpiteistä;

d. mallin elinkaaren kaikissa vaiheissa toteutetut tekniset ja organisatoriset toimenpiteet, jotka ovat joko edistäneet tai varmistaneet, ettei mallissa ole henkilötietoja;

e. asiakirjat, jotka osoittavat tekoälymallin teoreettisen kestävyyden torjua uudelleentunnistamisen teniikoita, sekä hallintatoimet, jotka on suunniteltu rajoittamaan tai

arvioimaan tärkeimpien hyökkäysten (regurgitaatio, joukkoon kuulumiseen perustuvaa päättely, eksfiltraatio jne.) onnistumista ja vaikutusta. Niitä voivat olla erityisesti seuraavat: I) koulutusdatan määrän ja parametrien lukumäärän välinen suhde mallissa, mukaan lukien analyysi sen vaikutuksesta malliin[38]; ii) uudelleentunnistamisen todennäköisyyden mittarit sillä hetkellä uusimman kehityksen perusteella; iii) raportit siitä, miten malli on testattu (kuka, milloin, miten ja missä laajuudessa) ja iv) testien tulokset;

f. mallin käyttöön ottavalle rekisterinpitäjälle tai rekisterinpitäjille ja/tai rekisteröidyille toimitetut asiakirjat, erityisesti asiakirjat, jotka liittyvät toteutettuihin toimenpiteisiin tunnistamisen todennäköisyyden vähentämiseksi, ja koskien mahdollisia jäännösriskejä.

3.3 Oikeutetun edun tarkoituksenmukaisuus henkilötietojen käsittelyn oikeusperusteena tekoälymallien kehittämisen ja käyttöönoton yhteydessä

59. Vastatakseen pyynnön toiseen ja kolmanteen kysymykseen tietosuojaneuvosto esittää ensin yleisiä huomioita joistakin tärkeistä näkökohdista, jotka valvontaviranomaisten olisi otettava huomioon käsittelyn oikeusperusteesta riippumatta arvioidessaan, miten rekisterinpitäjät voivat osoittaa noudattavansa yleistä tietosuoja-asetusta tekoälymallien yhteydessä. Tietosuojaneuvosto tarkastelee tämän jälkeen yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohtaan perustuvaa henkilötietojen käsittelyä koskevien ohjeiden (Guidelines 1/2024)[39] kolmea vaihetta, joita tarvitaan oikeutetun edun arvioinnissa tekoälymallien kehittämisen ja käyttöönoton yhteydessä.

3.3.1 Yleisiä huomioita

60. Tietosuojaneuvosto muistuttaa, että yleisessä tietosuoja-asetuksessa ei vahvisteta minkäänlaista hierarkiaa yleisen tietosuoja-asetuksen 6 artiklan 1 kohdassa säädettyjen eri oikeusperustojen välillä[40].

61. Yleisen tietosuoja-asetuksen 5 artiklassa vahvistetaan henkilötietojen käsittelyä koskevat periaatteet. Niistä tietosuojaneuvosto ottaa esiin ne, jotka ovat tämän lausunnon kannalta merkittäviä ja jotka valvontaviranomaisten olisi ainakin otettava huomioon arvioidessaan erityisiä tekoälymalleja, sekä tietosuoja-asetuksen muiden säännösten olennaisimmat vaatimukset, ottaen huomioon tämän lausunnon soveltamisala.

62. **Osoitusvelvollisuuden periaate** (yleisen tietosuoja-asetuksen 5 artiklan 2 kohta) – Tämän periaatteen mukaan rekisterinpitäjä vastaa yleisen tietosuoja-asetuksen noudattamisesta ja voi osoittaa sen noudattamisen. Tältä osin tekoälymallin kehittämisen tai käyttöönoton yhteydessä henkilötietoja käsittelevien osapuolten rooleja ja vastuita olisi arvioitava ennen käsittelyä, jotta voidaan määritellä rekisterinpitäjien tai yhteisrekisterinpitäjien sekä (mahdollisten) henkilötietojen käsittelijöiden velvollisuudet alusta alkaen.

63. **Lainmukaisuuden, kohtuullisuuden ja läpinäkyvyyden periaatteet** (yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan a alakohta) – Kun tietosuojaneuvosto arvioi käsittelyn lainmukaisuutta

[38] Ricciato F., *A Cautionary Reflection on (Pseudo-)Synthetic Data from Deep Learning on Personal Data*, Privacy in Statistical Databases -konferenssi (PSD 2024), Antibes, Ranska, syyskuu 2024, diat osoitteessa https://cros.ec.europa.eu/system/files/2024-10/20240926_PSD2024_Ricciato_v6_1.pdf ja Belkin M., Hsu D., Ma S., & Mandal S. (2019), *Reconciling modern machine-learning practice and the classical bias-variance trade-off*. Proceedings of the National Academy of Sciences, 24 July 2019, 116(32) 15849-15854, osoitteessa https://www.pnas.org/doi/10.1073/pnas.1903070116

[39] Ks. Euroopan tietosuojaneuvoston Guidelines 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024.

[40] Edellä mainittu 1 kohta.

tekoälymallien yhteydessä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan perusteella, se pitää hyödyllisenä erottaa eri vaiheet henkilötietojen käsittelyssä[41]. Oikeudenmukaisuuden periaate, joka liittyy läheisesti läpinäkyvyyden periaatteeseen, edellyttää, että henkilötietoja ei käsitellä epäoikeudenmukaisin menetelmin tai petoksen avulla tai "*rekisteröidyn kannalta perusteettoman haitallisella, lainvastaisen syrjivällä, odottamattomalla tai harhaanjohtavalla tavalla*"[42]. Kun otetaan huomioon asiaan liittyvien teknologioiden monimutkaisuus, tiedot henkilötietojen käsittelystä tekoälymalleissa olisi annettava helposti saatavilla olevalla, ymmärrettävällä ja käyttäjäystävällisellä tavalla[43]. Henkilötietojen käsittelyn avoimuuteen sisältyy erityisesti yleisen tietosuoja-asetuksen 12–14 artiklassa säädettyjen tiedonantovelvoitteiden noudattaminen [44] , mikä automaattisen päätöksenteon tapauksessa, mukaan lukien profilointi, edellyttää myös merkitseviä tietoja asiaan liittyvästä logiikasta sekä käsittelyn merkittävyydestä ja nähtävissä olevista seurauksista rekisteröidylle[45]. Kun otetaan huomioon, että tekoälymallien kehittämisvaiheissa saatetaan kerätä suuria määriä tietoja yleisesti saatavilla olevista lähteistä (esimerkiksi verkkosivujen haravointia hyödyntävien tekniikoiden avulla), yleisen tietosuoja-asetuksen 14 artiklan 5 kohdan b alakohdassa säädettyyn poikkeukseen turvautuminen on rajoitettu tiukasti siihen, että kyseisen säännöksen vaatimukset täyttyvät täysin[46].

64. **Rajoittamisen ja tietojen minimoinnin periaatteet** (yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan b ja c alakohta) – Tietojen minimoinnin periaatteen mukaisesti tekoälymallien kehittäminen ja käyttöönotto edellyttävät, että henkilötiedot ovat tarkoitukseen nähden asianmukaisia, olennaisia ja tarpeellisia. Tähän voi sisältyä henkilötietojen käsittely mahdollisten vinoumien ja virheiden riskien välttämiseksi, kun tämä on selkeästi ja nimenomaisesti yksilöity tarkoituksen yhteydessä ja henkilötiedot ovat välttämättömiä tätä tarkoitusta varten (esim. niitä ei voida saavuttaa tehokkaasti käsittelemällä muita tietoja, kuten synteettistä tai anonymisoitua dataa)[47]. Tietosuojatyöryhmä on jo korostanut, että *tietojenkeruun tarkoitus on yksilöitävä selkeästi ja täsmällisesti*[48]. Arvioitaessa, onko kyseinen tarkoitus laillinen, erityinen ja nimenomainen ja onko käsittely tietojen minimoinnin periaatteen mukaista, olisi ensin yksilöitävä kyseessä oleva käsittelytoimi. Erityisesti kehitys- tai käyttöönottovaiheen eri vaiheet voivat muodostua samoista tai erilaisista käsittelytoimista, ja niihin voi sisältyä peräkkäisiä rekisterinpitäjiä tai yhteisrekisterinpitäjiä. Joissakin tapauksissa on mahdollista määrittää tekoälymallin käyttöönoton tarkoitus varhaisessa kehitysvaiheessa. Vaikka näin ei olisikaan, käyttöönoton jonkinlaisen kontekstin olisi jo oltava selkeä, ja siksi olisi harkittava, miten tämä konteksti vaikuttaa kehityksen tarkoitukseen. Tarkastellessaan käsittelyn tarkoitusta tietyssä kehitysvaiheessa

[41] Euroopan tietosuojaneuvoston 23. toukokuuta 2024 hyväksymän ChatGPT-työryhmän työtä käsittelevän raportin 14 kohta.

[42] EDPB:n raportti ChatGPT-työryhmän työstä, hyväksytty 23. toukokuuta 2024, 23 kohta; EDPB:n ohjeet 4/2019 25 artiklan mukaisesta sisäänrakennetusta ja oletusarvoisesta tietosuojasta, versio 2.0, annettu 20. lokakuuta 2020, 69 kohta; tietosuojatyöryhmän asetuksen 2016/679 mukaista läpinäkyvyyttä koskevat suuntaviivat, viimeksi tarkistettu ja hyväksytty 11. huhtikuuta 2018, EDPB hyväksynyt 25. toukokuuta 2018, 2 kohta.

[43] Tietosuojatyöryhmän asetuksen (EU) N:o 2016/679 mukaista läpinäkyvyyttä koskevat suuntaviivat, viimeksi tarkistettu ja hyväksytty 11. huhtikuuta 2018, EDPB hyväksynyt 25. toukokuuta 2018, 5 kohta.

[44] Ks. yleisen tietosuoja-asetuksen johdanto-osan 39 kappale, jonka mukaan ”*[L]uonnollisille henkilöille olisi oltava läpinäkyvää, miten heitä koskevia henkilötietoja kerätään ja käytetään ja niihin tutustutaan tai niitä käsitellään muulla tavoin sekä selvillä siitä, missä määrin henkilötietoja käsitellään tai on määrä käsitellä*”.

[45] Yleisen tietosuoja-asetuksen 13 artiklan 2 kohdan f alakohta ja 14 artiklan 2 kohdan g alakohta.

[46] EDPB:n raportti ChatGPT-työryhmän työstä, hyväksytty 23. toukokuuta 2024, 27 kohta.

[47] Lisäksi tekoälysäädöksen 10 artiklan 5 kohdassa säädetään erityisistä säännöistä, jotka koskevat suuririskisiin tekoälyjärjestelmiin liittyvien erityisten henkilötietoryhmien käsittelyä vinouumien havaitsemisen ja korjaamisen varmistamiseksi.

[48] Tietosuojatyöryhmän lausunto Opinion 03/2013 on purpose limitation (WP203), s. 15–16.

valvontaviranomaisten olisi odotettava rekisterinpitäjältä (rekisterinpitäjiltä) tiettyä yksityiskohtaisuutta ja selitystä siitä, miten nämä yksityiskohdat vaikuttavat käsittelyn tarkoitukseen. Tähän voi sisältyä esimerkiksi tietoja kehitetyn tekoälymallin tyypistä, sen odotetuista toiminnoista ja muusta asiaankuuluvasta kontekstista, joka on jo kyseisessä vaiheessa tiedossa. Käyttöönoton kontekstiin voisi kuulua myös esimerkiksi se, kehitetäänkö mallia sisäistä käyttöönottoa varten, aikooko rekisterinpitäjä myydä tai jakaa mallia kolmansille osapuolille sen kehittämisen jälkeen, mukaan lukien se, onko malli ensisijaisesti tarkoitettu käytettäväksi tutkimustarkoituksiin tai kaupallisiin tarkoituksiin.

65. **Rekisteröidyn oikeudet** (yleisen tietosuoja-asetuksen III luku) – Vaikka valvontaviranomaisten on varmistettava, että kaikkia rekisteröidyn oikeuksia kunnioitetaan, kun rekisterinpitäjät kehittävät ja ottavat käyttöön tekoälymalleja, tietosuojaneuvosto muistuttaa, että aina kun rekisterinpitäjä käyttää oikeutettua etua oikeusperusteena, sovelletaan yleisen tietosuoja-asetuksen 21 artiklan mukaista vastustamisoikeutta, ja se olisi varmistettava[49].

3.3.2 Huomioita oikeutetun edun arvioinnin kolmesta vaiheesta tekoälymallien kehittämisen ja käyttöönoton yhteydessä

66. Määrittääkseen, voiko tietty henkilötietojen käsittely perustua tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohtaan, valvontaviranomaisten olisi tarkistettava, että rekisterinpitäjät ovat huolellisesti arvioineet ja dokumentoineet, täyttyvätkö seuraavat kolme kumulatiivista ehtoa: i) rekisterinpitäjän tai kolmannen osapuolen oikeutetun edun toteuttaminen, ii) käsittely on tarpeen oikeutetun edun toteuttamiseksi, ja iii) rekisteröityjen edut tai perusoikeudet ja -vapaudet eivät syrjäytä oikeutettua etua[50].

3.3.2.1 Ensimmäinen vaihe – Rekisterinpitäjän tai kolmannen osapuolen oikeutetun edun tavoittelu

67. Etu on laajempi intressi tai hyöty, joka rekisterinpitäjällä tai kolmannella osapuolella voi olla tiettyyn käsittelytoimeen osallistumisesta [51] . Vaikka yleisessä tietosuoja-asetuksessa ja unionin

[49] Yleisen tietosuoja-asetuksen 21 artiklan mukaan, jos rekisteröity vastustaa erityiseen tilanteeseensa liittyvällä perusteella häntä koskevien henkilötietojen käsittelyä, rekisterinpitäjä ei saa enää käsitellä henkilötietoja, paitsi jos rekisterinpitäjä voi osoittaa, että käsittelyyn on olemassa huomattavan tärkeä ja perusteltu syy, joka syrjäyttää rekisteröidyn edut, oikeudet ja vapaudet tai jos se on tarpeen oikeusvaateen laatimiseksi, esittämiseksi tai puolustamiseksi. Näin ollen valvontaviranomaisten on otettava huomioon kaksi seikkaa: kykeneekö rekisterinpitäjä osoittamaan tällaisen huomattavan tärkeän ja perustellun syyn, joka syrjäyttää rekisteröidyn edut, ja voidaanko vastustamisoikeutta käyttää.

[50] Unionin tuomioistuimen tuomio 4. heinäkuuta 2023 asiassa C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), 106 kohta; unionin tuomioistuimen tuomio 11. joulukuuta 2019 asiassa C-708/18, *Asociaţia de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), 40 kohta. Ks. myös tietosuojaneuvoston Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, Version 1.0 adopted on 8 October 2024, paragraph 12 and ff1/2024 (ohjeet henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 12 kohta ja sitä seuraavat kohdat). Kuten näissä ohjeissa muistutetaan, tämä *arviointi olisi tehtävä käsittelyn alussa tietosuojavastaavan (jos sellainen on nimetty) osallistuessa siihen, ja rekisterinpitäjän olisi dokumentoitava se yleisen tietosuoja-asetuksen 5 artiklan 2 kohdassa säädetyn osoitusvelvollisuusperiaatteen mukaisesti*.

[51] Euroopan tietosuojaneuvoston Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, versio 1.0, annettu 8. lokakuuta 2024, 14 kohta (ohjeet henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella).

tuomioistuimessa on tunnustettu useita etuja oikeutetuiksi[52], tietyn edun oikeutuksen arvioinnin olisi perustuttava tapauskohtaiseen analyysiin.

68. Kuten Euroopan tietosuojaneuvosto muistuttaa oikeutettua etua koskevissa ohjeissaan [53], etua voidaan pitää oikeutettuna, jos seuraavat kolme kumulatiivista kriteeriä täyttyvät:

    a. intressi on laillinen[54]

    b. intressi on ilmaistu selkeästi ja täsmällisesti ja

    c. Intressi on todellinen ja läsnä, ei spekulatiivinen.

69. Jollei oikeutetun edun arvioinnin edellyttämistä kahdesta muusta vaiheesta muuta johdu, seuraavat esimerkit voivat muodostaa tekoälymallien yhteydessä oikeutetun edun: i) keskustelevan palvelun järjestelmän kehittäminen käyttäjien avustamiseksi, ii) tekoälyjärjestelmän kehittäminen vilpillisen sisällön tai käyttäytymisen havaitsemiseksi ja iii) uhkien havaitsemisen parantaminen tietojärjestelmässä.

3.3.2.2 Toinen vaihe – analyysi käsittelyn tarpeellisuudesta oikeutetun edun toteuttamiseksi

70. Arvioinnin toisessa vaiheessa määritellään, onko henkilötietojen käsittely tarpeen oikeutetun edun (oikeutettujen etujen) saavuttamiseksi[55] (ns. tarpeellisuustesti).

71. Yleisen tietosuoja-asetuksen johdanto-osan 39 kappaleessa todetaan, että ”[H]enkilötietoja olisi käsiteltävä vain jos käsittelyn tarkoitusta ei voida kohtuullisesti toteuttaa muilla keinoin”. Unionin tuomioistuimen ja Euroopan tietosuojaneuvoston aiempien ohjeiden mukaan käsittelyn tarpeellisuutta koskevaa edellytystä olisi tarkasteltava rekisteröityjen perusoikeuksien ja -vapauksien valossa ja yhdessä yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan c alakohdassa vahvistetun tietojen minimoinnin periaatteen kanssa[56].

---

[52] Euroopan tietosuojaneuvoston Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, versio 1.0, annettu 8. lokakuuta 2024, 16 kohta (ohjeet henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella).

[53] Euroopan tietosuojaneuvoston Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, versio 1.0, annettu 8. lokakuuta 2024, 17 kohta (ohjeet henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella).

[54] Unionin tuomioistuimen tuomio 4. lokakuuta 2024 asiassa C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), 49 kohta, jossa unionin tuomioistuin painottaa, että oikeutettu etu ei voi olla lainvastainen. Tältä osin Euroopan tietosuojaneuvosto korostaa, että tapauksen mukaan lainsäädännölliset puitteet olisi otettava huomioon arvioitaessa tietyn edun laillisuutta. Ks. esimerkiksi: Digitaalisten palvelujen sisämarkkinoista ja direktiivin 2000/31/EY muuttamisesta 19 päivänä lokakuuta 2022 annetun Euroopan parlamentin ja neuvoston asetuksen (EU) 2022/2065 (’digipalvelusäädös’) 26 artiklan 3 kohta ja 28 artikla kielletystä alaikäisille kohdennetusta mainonnasta; tekoälysäädöksen 5 artiklan 1 ja 2 kohta kielletyistä tekoälyyn liittyvistä käytännöistä (manipulatiiviset käytännöt ja tietoisuuskynnyksen alittaminen); teollis- ja tekijänoikeuksia loukkaavasta käsittelystä sekä tekijänoikeudesta ja lähioikeuksista digitaalisilla sisämarkkinoilla annetun direktiivin (EU) 2019/790 säännökset.

[55] Euroopan tietosuojaneuvoston ohjeet henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 28–30 kohta.

[56] Unionin tuomioistuimen tuomio 4. heinäkuuta 2023 asiassa C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), 108 ja 109 kohta, jossa viitataan myös unionin tuomioistuimen tuomioon 11. joulukuuta 2019 asiassa C-708/18, *Asociaţia de Proprietari bloc M5A-Scara* (ECLI:EU:C:2019:1064), 48 kohta; unionin tuomioistuimen tuomio 9. marraskuuta 2010 yhdistetyissä asioissa C-92/09 ja C-93/09, Volker und Markus Schecke (ECLI:EU:C:2010:662), 85 ja 86 kohta; unionin tuomioistuimen tuomio 22. kesäkuuta 2021 asiassa C-439/19, *Latvijas Republikas Saeima* (ECLI:EU:C:2021:504), 98, 109, 110 ja 113 kohta. Ks. Esimerkiksi myös: Euroopan tietosuojaneuvoston ohjeet 3/2019 henkilötietojen käsittelystä videolaitteilla, versio 2.0, annettu

72. Unionin tuomioistuimen mainitsemassa menetelmässä otetaan huomioon käsittelyn asiayhteys sekä vaikutukset rekisterinpitäjään ja rekisteröityihin. Tarpeellisuuden arviointiin sisältyy näin ollen kaksi seikkaa: (i) mahdollistaako käsittelytoiminta kyseisen tarkoituksen toteuttamisen [57] ja (ii) onko vähemmän puuttuvaa tapaa toteuttaa tätä tarkoitusta[58].

73. Esimerkiksi tekoälymalliin liittyvien henkilötietojen aiottua määrää on tapauksen mukaan arvioitava ottaen huomioon vähemmän puuttuvat vaihtoehdot, jotka voivat kohtuudella olla käytettävissä ja joilla voidaan yhtä tehokkaasti saavuttaa tavoitellun oikeutetun edun tarkoitus. Jos tarkoituksen toteuttaminen on mahdollista myös sellaisen tekoälymallin avulla, joka ei edellytä henkilötietojen käsittelyä, henkilötietojen käsittelyä ei olisi pidettävä tarpeellisena. Tämä on erityisen tärkeää tekoälymallien kehittämisen kannalta. Arvioidessaan, täyttyykö välttämättömyyden edellytys, valvontaviranomaisten olisi kiinnitettävä erityistä huomiota käsiteltävien henkilötietojen määrään ja siihen, onko se oikeassa suhteessa kyseessä olevaan oikeutettuun etuun, myös tietojen minimointiperiaatteen valossa.

74. Tarpeellisuuden arvioinnissa olisi myös otettava huomioon henkilötietojen suunnitellun käsittelyn laajempi asiayhteys. Rekisteröityjen perusoikeuksia ja -vapauksia vähemmän loukkaavien keinojen olemassaolo voi vaihdella sen mukaan, onko rekisterinpitäjällä suora suhde rekisteröityihin (ensimmäisen osapuolen tiedot) vai ei (kolmannen osapuolen tiedot). Euroopan unionin tuomioistuin esitti joitakin näkökohtia, jotka on otettava huomioon analysoitaessa, onko ensimmäisen osapuolen tietojen käsittely tarpeen oikeutetun edun (oikeutettujen etujen) toteuttamiseksi (vaikkakin kyseisten tietojen luovuttamisen yhteydessä kolmansille osapuolille)[59].

75. Teknisten suojatoimien toteuttaminen henkilötietojen suojaamiseksi voi myös edistää tarpeellisuustestin noudattamista. Tämä voisi sisältää 3.2.2 kohdassa yksilöityjen toimenpiteiden

---

29. tammikuuta 2020, kohdat 24–26 ja 73; Euroopan tietosuojaneuvoston ohjeet 2/2019 yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan b alakohdan perusteella tapahtuvasta henkilötietojen käsittelystä rekisteröidyille tarjottavien verkkopalvelujen yhteydessä, versio 2.0, annettu 8. lokakuuta 2019, kohdat 23–25; Euroopan tietosuojaneuvoston lausunto 11/2024 kasvojentunnistuksen käytöstä lentoasemien matkustajavirtojen sujuvoittamiseksi, versio 1.1, annettu 23. toukokuuta 2024, kohta 27.

[57] Ks. Euroopan unionin tuomioistuimen tuomio, 16 päivänä joulukuuta 2008, asiassa C-524/06, *Heinz Huber v. Bundesrepublik Deutschland* (ECLI:EU:C:2008:724), 66 kohta. Ks. myös samassa asiassa julkisasiamies Poiares Maduron ratkaisuehdotus asiassa C-524/06, *Heinz Huber vastaan Saksan liittotasavalta (*ECLI:EU:C:2008:194), 16 kohta, jossa todetaan seuraavaa: '- - *ratkaisevaa on tehokkuus, ja tämän seikan arviointi on kansallisen tuomioistuimen tehtävä. Sen on kysyttävä, onko olemassa muita tietojenkäsittelytapoja, joiden avulla maahanmuuttoviranomaiset voivat varmistaa oleskeluoikeudellista asemaa koskevien sääntöjen noudattamisen. Jos se vastaa tähän kysymykseen myöntävästi, unionin kansalaisten tietojen keskitettyä tallentamista ja käsittelyä olisi pidettävä lainvastaisena. Vaihtoehtoisen järjestelmän ei välttämättä tarvitse olla kaikkein tehokkain tai sopivin; riittää, että se kykenee suoriutumaan tehtävästään. Vaikka keskusrekisteri olisikin siis tehokkaampi, kätevämpi tai helppokäyttöisempi kuin muut vaihtoehdot (kuten hajautetut, paikalliset rekisterit), viimeksi mainitut ovat selvästi suositeltavia, jos niiden avulla voidaan osoittaa unionin kansalaisten oleskeluoikeudellinen asema*.

[58] Ks. unionin tuomioistuimen tuomio, 27 päivänä syyskuuta 2017, asiassa C-73/16, *Peter Puškár* (ECLI:EU:C:2017:725), 113 kohta: "*Ennakkoratkaisua pyytäneen tuomioistuimen on siis tarkistettava, onko niin, että riidanalaisen luettelon laatiminen ja rekisteröityjen nimien merkitseminen luetteloon soveltuvat niillä tavoiteltujen päämäärien saavuttamiseksi ja ettei näiden päämäärien saavuttamiseksi ole olemassa lievempiä keinoja*"; ks. myös esimerkiksi julkisasiamies Rantosin ratkaisuehdotus asiassa C-252/21, *Meta v. Bundeskartellamt*, ECLI:EU:C:2022:704, 61 kohta, jonka mukaan *"[...] Käsittelyn ja tavoitellun edun välillä on siksi oltava läheinen yhteys silloin, kun ei ole vaihtoehtoja, joissa henkilötietoja suojattaisiin paremmin, sillä ei riitä, että käsittelystä on hyötyä ainoastaan rekisterinpitäjälle."*

[59]Unionin tuomioistuimen tuomio, 4 päivänä lokakuuta 2024, asiassa C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), 51–53 kohta.

kaltaisia toimenpiteistä siten, ettei anonymisointiin pyritä, mutta vähennetään rekisteröityjen tunnistamisen helppoutta. Tietosuojaneuvosto toteaa, että jotkin näistä toimenpiteistä, vaikka niitä ei vaadita yleisen tietosuoja-asetuksen noudattamiseksi, voivat olla lisäsuojatoimia, kuten 3.3.2.3 jakson kohdassa Lieventävät toimenpiteet tarkemmin analysoidaan[60].

3.3.2.3 Kolmas vaihe – Tasapainotesti

76. Oikeutetun edun arvioinnin kolmas vaihe on **tasapainottaminen** (myös **tasapainotesti** tässä asiakirjassa)[61]. Tässä vaiheessa yksilöidään ja kuvataan kyseessä olevat erilaiset vastakkaiset oikeudet ja edut[62] eli yhtäältä rekisteröityjen edut, perusoikeudet ja -vapaudet ja toisaalta rekisterinpitäjän tai kolmannen osapuolen edut. Tapauksen erityisolosuhteiden olisi tällöin katsottava osoittavan, että oikeutettu etu on asianmukainen oikeusperuste kyseisille käsittelytoimille[63].

Rekisteröityjen edut, perusoikeudet ja -vapaudet

77. Yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdassa säädetään, että arvioidessaan eri osatekijöitä tasapainotestin yhteydessä rekisterinpitäjän on otettava huomioon rekisteröityjen edut, perusoikeudet ja -vapaudet. Rekisteröityjen etuja ovat ne, joihin kyseinen käsittely voi vaikuttaa. Tekoälymallin kehittämisvaiheen yhteydessä niihin voi sisältyä muun muassa, mutta ei pelkästään, etuja, jotka liittyvät itsemääräämisoikeuteen ja määräysvallan säilyttämiseen omien henkilötietojen suhteen (esim. mallin kehittämistä varten kerättyjen osalta). Tekoälymallin käyttöönoton yhteydessä rekisteröityjen etuja voivat olla muun muassa määräysvallan säilyttäminen omiin henkilötietoihin (esim. mallin käyttöönoton jälkeen käsiteltävät tiedot), taloudelliset edut (esim. jos rekisteröity käyttää tekoälymallia tulojen tuottamiseen tai yksilö käyttää sitä ammatillisen toimintansa yhteydessä), henkilökohtainen hyöty (esim. kun tekoälymallia käytetään parantamaan tiettyjen palvelujen saatavuutta) tai sosioekonomiset edut (esim. kun tekoälymalli mahdollistaa paremman terveydenhuollon tai tukee perusoikeuden, kuten koulutukseen pääsyn, käyttöä)[64].

78. Mitä tarkemmin etu määritellään käsittelyn aiotun tarkoituksen perusteella, sitä paremmin sen avulla on mahdollista ymmärtää selkeästi, mitkä hyödyt ja riskit on otettava huomioon tasapainotestissä.

79. Rekisteröityjen perusoikeuksien ja -vapauksien osalta tekoälymallien kehittäminen ja käyttöönotto voivat aiheuttaa vakavia riskejä EU:n perusoikeuskirjassa suojattuihin oikeuksiin, kuten yksityis- ja perhe-elämän kunnioittamiseen (EU:n perusoikeuskirjan 7 artikla) ja henkilötietojen suojaan (EU:n perusoikeuskirjan 8 artikla). Näitä riskejä voi esiintyä kehittämisvaiheessa esimerkiksi silloin, kun henkilötietoja haravoidaan vastoin rekisteröityjen toiveita tai heidän tietämättään. Näitä riskejä voi esiintyä myös käyttöönottovaiheessa, esimerkiksi silloin, kun mallissa (tai sen osassa) käsitellään henkilötietoja tavalla, joka on vastoin rekisteröityjen oikeuksia, tai kun on mahdollista päätellä vahingossa tai hyökkäysten perusteella (esim. joukkoon kuulumiseen perustuva päättely, tietojen poiminnan tai mallin invertointi), mitä henkilötietoja koulutustietokanta sisältää. Tällaiset tilanteet

[60] Ks. Euroopan tietosuojaneuvoston ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 57 kohta.

[61] Ks. Euroopan tietosuojaneuvoston ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 31–60 kohta (EDPB Guidelines 1/2024 on processing of personal data).

[62] Ks. Euroopan tietosuojaneuvoston ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 32 kohta.

[63] Ks. Euroopan tietosuojaneuvoston ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 32 kohta, ks. myös Euroopan unionin tuomioistuimen tuomio 4 päivänä heinäkuuta 2023, asiassa C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), 110 kohta.

[64] Ks. Euroopan tietosuojaneuvoston ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 38 kohta.

ovat riski niiden rekisteröityjen yksityisyydelle, joiden tiedot saattavat näkyä tekoälyjärjestelmän käyttöönottovaiheessa (esimerkiksi maineriski, identiteettivarkaus tai -petos, turvallisuusriski tietojen luonteen mukaan).

80. Käsiteltävän tapauksen mukaan myös muihin perusoikeuksiin voi kohdistua riskejä. Esimerkiksi tekoälymallien tekemä laajamittainen ja rajoittamaton tiedonkeruu kehittämisvaiheessa voi tuntua rekisteröidyistä valvonnalta, erityisesti kun otetaan huomioon vaikeudet estää julkisen datan haravointi. Tämä voi johtaa henkilöiden itsesensuuriin, ja riskinä on sananvapauden heikkeneminen (EU:n perusoikeuskirjan 11 artikla). Myös käyttöönottovaiheessa on sananvapauteen kohdistuvia riskejä, kun tekoälymalleja käytetään estämään sisällön julkaiseminen rekisteröidyiltä. Lisäksi tekoälymalli, joka suosittelee haavoittuvassa asemassa oleville henkilöille sopimatonta sisältöä, voi aiheuttaa riskin heidän mielenterveydelleen (EU:n perusoikeuskirjan 3 artiklan 1 kohta). Joittakin tapauksissa tekoälymallien käyttöönotto voi myös vaikuttaa haitallisesti yksilön oikeuteen tehdä työtä (EU:n perusoikeuskirjan 15 artikla), esimerkiksi kun työhakemusten esikarsinta tehdään tekoälymallin avulla. Vastaavasti tekoälymalli voi aiheuttaa riskejä syrjimättömyydelle (EU:n perusoikeuskirjan 21 artikla), jos yksilöitä syrjitään tiettyjen henkilökohtaisten ominaisuuksien (kuten kansalaisuuden tai sukupuolen) perusteella. Lisäksi tekoälymallien käyttöönotto voi aiheuttaa riskejä yksilön turvallisuudelle (esim. jos tekoälymallia käytetään vihamielisessä tarkoituksessa) sekä riskejä fyysiselle ja henkiselle koskemattomuudelle[65].

81. Tekoälymallien käyttöönotto voi myös vaikuttaa myönteisesti tiettyihin perusoikeuksiin. Malli voi esimerkiksi tukea henkilön oikeutta henkiseen koskemattomuuteen (perusoikeuskirjan 3 artikla) tilanteessa, jossa tekoälymallia käytetään haitallisen verkkosisällön tunnistamiseen, tai malli voi helpottaa tiettyjen keskeisten palvelujen saatavuutta tai perusoikeuksien käyttöä, kuten tiedonsaantia (EU:n perusoikeuskirjan 11 artikla) tai pääsyä koulutukseen (EU:n perusoikeuskirjan 14 artikla).

Käsittelyn vaikutus rekisteröityihin

82. Tekoälymallien kehittämisen ja käyttöönoton aikana tapahtuva henkilötietojen käsittely voi vaikuttaa rekisteröityihin eri tavoin, ja se voi olla myönteistä tai kielteistä[66]. Esimerkiksi jos käsittelytoimesta on rekisteröidylle hyötyä, se voidaan ottaa huomioon tasapainotestissä. Koska tällaisen hyödyn olemassaolo voi johtaa siihen, että valvontaviranomainen katsoo, että rekisteröityjen edut, perusoikeudet ja -vapaudet eivät syrjäytä rekisterinpitäjän tai kolmannen osapuolen etuja, tällainen päätelmä voi perustua ainoastaan tapauskohtaiseen analyysiin, jossa otetaan huomioon kaikki asianmukaiset tekijät.

83. Käsittelyn vaikutukseen rekisteröityihin voivat vaikuttaa i) mallien käsittelemien tietojen luonne, ii) käsittelyn konteksti ja iii) käsittelyn mahdolliset muut seuraukset[67].

84. Käsiteltävien tietojen **luonteesta** on muistettava, että – lukuun ottamatta erityisiä henkilötietoryhmiä ja rikostuomioihin ja rikoksiin liittyviä tietoja, jotka nauttivat tietosuoja-asetuksen 9 ja 10 artiklan mukaista lisäsuojaa – joidenkin muiden henkilötietoryhmien käsittelystä voi aiheutua merkittäviä seurauksia rekisteröidyille. Tässä yhteydessä olisi otettava huomioon, että sellaisten tietyntyyppisten henkilötietojen käsittely tekoälymallin kehittämistä ja käyttöönottoa varten, jossa paljastuu erittäin yksityisiä tietoja (esim. rahoitustietoja tai sijaintitietoja), voi vaikuttaa vakavasti rekisteröityihin.

[65] Ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, hyväksytty 8. lokakuuta 2024, 46 kohta.
[66] Ks. Euroopan tietosuojaneuvoston ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 39 kohta.
[67] Ks. Euroopan tietosuojaneuvoston ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 32 kohta.

Tällaisen käsittelyn seuraukset käyttöönottovaiheessa voivat olla rekisteröidyille esimerkiksi taloudellisia (esim. syrjintä työelämään liittyvissä asioissa) ja/tai liittyä maineeseen (esim. kunnianloukkaus).

85. **Käsittelyn yhteydessä** on ensin yksilöitävä tekijät, jotka voivat aiheuttaa riskejä rekisteröidyille (esim. tapa, jolla malli on kehitetty, tapa, jolla malli voidaan ottaa käyttöön ja/tai ovatko henkilötietojen suojaamiseen käytetyt turvatoimet asianmukaisia). Mallin luonne ja suunnitellut operatiiviset käyttötarkoitukset ovat keskeisessä asemassa tällaisten mahdollisten syiden tunnistamisessa.

86. On myös tarpeen arvioida näiden riskien vakavuutta rekisteröityjen kannalta. Huomioon voidaan ottaa muun muassa se, miten henkilötietoja käsitellään (esim. niiden yhdistäminen muihin tietoaineistoihin), mikä on käsittelyn laajuus ja käsiteltävien henkilötietojen määrä [68] (esim. datan kokonaismäärä, tietojen määrä rekisteröityä kohti, niiden rekisteröityjen määrä, joihin käsittely vaikuttaa) [69], rekisteröidyn asema (esim. lapset tai muut haavoittuvassa asemassa olevat rekisteröidyt) ja heidän suhteensa rekisterinpitäjään (esim. rekisteröity on asiakas). Esimerkiksi verkkosivujen haravointi kehittämisvaiheessa voi johtaa – ilman riittäviä suojatoimia – merkittäviin yksilöihin kohdistuviin vaikutuksiin kerättyjen tietojen suuren määrän, rekisteröityjen suuren määrän ja henkilötietojen rajoittamattoman tietojenkeruun vuoksi.

87. Myös käsittelyn mahdolliset **muut seuraukset** olisi otettava huomioon arvioitaessa käsittelyn vaikutusta rekisteröityihin. Valvontaviranomaisten olisi tehtävä arvioinnit tapauskohtaisesti ottaen huomioon kulloinkin kyseessä olevat erityiset tosiseikat.

88. Tällaisia seurauksia voivat olla muun muassa (mutta eivät ainoastaan) riskit rekisteröityjen perusoikeuksien loukkaamisesta, kuten edellisessä kohdassa [70] on kuvattu. Riskit voivat vaihdella todennäköisyydeltään ja vakavuudeltaan. Ne voivat johtua henkilötietojen käsittelystä ja aiheuttaa fyysistä, aineellista tai aineetonta vahinkoa, erityisesti jos käsittely johtaa syrjintään[71].

89. Jos tekoälymallin käyttöönotto edellyttää sekä i) sellaisten rekisteröityjen henkilötietojen käsittelyä, joiden henkilötiedot sisältyvät kehittämisvaiheessa käytettyyn tietoaineistoon, että ii) sellaisten rekisteröityjen henkilötietojen käsittelyä, joiden henkilötietoja käsitellään käyttöönottovaiheessa, valvontaviranomaisten olisi rekisterinpitäjän suorittamaa tasapainotestiä tarkistaessaan erotettava toisistaan ja otettava huomioon ne riskit, jotka vaikuttavat kummankin rekisteröityjen ryhmän etuihin, oikeuksiin ja vapauksiin.

90. **Lopuksi käsittelyn mahdollisten myöhempien seurausten analyysissä olisi otettava huomioon myös näiden myöhempien seurausten toteutumisen todennäköisyys**. Tällaisen todennäköisyyden arvioinnissa olisi otettava huomioon käytössä olevat tekniset ja organisatoriset toimenpiteet sekä tapauksen erityisolosuhteet. Valvontaviranomaiset voivat esimerkiksi tarkastella, onko toteutettu toimenpiteitä, joilla vältetään tekoälymallin mahdollista väärinkäyttöä. Kun kyse on tekoälymalleista, joita voidaan käyttää eri tarkoituksissa, kuten generatiivisessa tekoälyssä, tähän voi sisältyä hallintatoimia, jotka rajoittavat mahdollisimman paljon niiden käyttöä haitallisiin käytäntöihin, kuten esimerkiksi syvävääreennösten luomiseen, disinformaatiota käyttäviin asiointibotteihin, tietojen

[68] Ks. Euroopan tietosuojaneuvoston ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 43 kohta.

[69] Unionin tuomioistuimen tuomio 4 päivänä heinäkuuta 2023 asiassa C-252/21 *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), 116 kohta.

[70] Ks. edellä kohta Rekisteröityjen edut, perusoikeudet ja -vapaudet.

[71] Ks. Euroopan tietosuojaneuvoston ohjeiden 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024 jakso 2.3. Lisää esimerkkejä on myös yleisen tietosuoja-asetuksen johdanto-osan 75 kappaleessa.

kalasteluun ja muunlaisiin petoksiin sekä manipuloiviin tekoälyihin ja tekoälyjärjestelmiin (erityisesti, kun ne ovat antropomorfisia tai antavat harhaanjohtavaa tietoa).

Rekisteröityjen kohtuulliset odotukset

91. Perustuu yleisen tietosuoja-asetuksen johdanto-osan 47 kappaleeseen *”[O]ikeutetun edun olemassaoloa on joka tapauksessa arvioitava huolellisesti; on arvioitava muun muassa, voiko rekisteröity kohtuudella odottaa henkilötietojen keruun ajankohtana ja sen yhteydessä, että henkilötietoja voidaan käsitellä tätä tarkoitusta varten. Etenkin rekisteröidyn edut ja perusoikeudet voisivat syrjäyttää rekisterinpitäjän edun, jos henkilötietoja käsitellään olosuhteissa, joissa rekisteröity ei voi kohtuudella odottaa jatkokäsittelyä*. *Etenkin rekisteröidyn edut ja perusoikeudet voisivat syrjäyttää rekisterinpitäjän edun, jos henkilötietoja käsitellään olosuhteissa, joissa rekisteröity ei voi kohtuudella odottaa jatkokäsittelyä”.*[72]

92. Kohtuulliset odotukset ovat avainasemassa tasapainotestissä, eikä vähiten tekoälymalleissa käytettävän teknologian monimutkaisuuden vuoksi ja siksi, että rekisteröityjen voi olla vaikea ymmärtää tekoälymallin erilaisia mahdollisia käyttötapoja ja siihen liittyvää tietojenkäsittelyä[73]. Tätä varten rekisteröidyille annettuja tietoja voidaan tarkastella sen arvioimiseksi, voivatko rekisteröidyt kohtuudella odottaa, että heidän henkilötietojaan käsitellään. Vaikka tietojen antamatta jättäminen voi vaikuttaa siihen, että rekisteröidyt eivät odota tiettyä käsittelyä, pelkästään yleisessä tietosuoja-asetuksessa säädettyjen läpinäkyvyysvaatimusten täyttäminen ei itsessään riitä siihen, että voidaan katsoa, että rekisteröidyt voivat kohtuudella odottaa tiettyä käsittelyä[74]. Lisäksi pelkästään se, että tekoälymallin kehittämisvaiheeseen liittyvät tiedot sisältyvät rekisterinpitäjän tietosuojapolitiikkaan, ei välttämättä tarkoita, että rekisteröidyt voivat kohtuudella odottaa, että näin tapahtuu. Valvontaviranomaisten olisi pikemminkin analysoitava tätä tapauksen erityisolosuhteiden ja kaikkien merkityksellisten tekijöiden perusteella.

93. Arvioitaessa rekisteröityjen kohtuullisia odotuksia tietojenkäsittelystä, joka tapahtuu kehittämisvaiheessa, on tärkeää viitata seikkoihin, jotka mainitaan Euroopan tietosuojaneuvoston ohjeissa oikeutetusta edusta[75]. Lisäksi tämän lausunnon aihepiirissä on tärkeää tarkastella käsittelyn laajempaa kontekstia. Tähän voi sisältyä muun muassa se, ovatko henkilötiedot olleet julkisesti saatavilla, rekisteröidyn ja rekisterinpitäjän välisen suhteen luonne (ja se, onko näiden kahden välillä olemassa yhteys), palvelun luonne, henkilötietojen keräämisen asiayhteys, lähde, josta tiedot on kerätty (esim. verkkosivusto tai palvelu, jolta henkilötiedot on kerätty, ja sen tarjoamaa yksityisyyden suojaa koskevat asetukset), mallin mahdollinen lisäkäyttö ja se, ovatko rekisteröidyt tosiasiallisesti selvillä siitä, että heidän henkilötietonsa ylipäätään ovat verkossa.

---

[72] Ks. myös unionin tuomioistuimen tuomio 4 päivänä heinäkuuta 2023 asiassa C-252/21 *Meta vastaan Bundeskartellamt* (ECLI:EU:C:2023:537), 112 kohta; unionin tuomioistuimen tuomio 11 päivänä joulukuuta 2019 asiassa C-708/18 *Asociaţia de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), 58 kohta; unionin tuomioistuimen tuomio 4 päivänä lokakuuta 2024 asiassa C-621/22 *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), 55 kohta.

[73] Esimerkiksi 4. heinäkuuta 2023 asiassa C-252/21, *Meta v. Bundeskartellamt* (ECLI:EU:C:2023:537), annetun tuomion 123 kohta, vaikka unionin tuomioistuin totesikin, että ”tuotteiden parantamisen” tavoitetta ei voida periaatteessa sulkea pois oikeutettuna etuna, se totesi myös, että ”*on kyseenalaista, että [...] tuotteiden parantamista koskeva tavoite voisi, kun otetaan huomioon tämän käsittelyn laajuus ja sillä käyttäjään oleva huomattava vaikutus sekä se, että käyttäjä ei voi kohtuudella odottaa [...] käsittelevän näitä tietoja, syrjäyttää tällaisen käyttäjän edut ja perusoikeudet erityisesti tilanteessa, jossa käyttäjä on lapsi*”.

[74] Ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, hyväksytty 8. lokakuuta 2024, 53 kohta.

[75] Ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 50–54 kohta.

94. Mallin kehittämisvaiheessa rekisteröityjen kohtuulliset odotukset voivat vaihdella sen mukaan, julkistavatko rekisteröidyt mallin kehittämiseksi käsiteltävät tiedot vai eivät. Lisäksi kohtuulliset odotukset voivat vaihdella sen mukaan, ovatko he toimittaneet tiedot suoraan rekisterinpitäjälle (esimerkiksi palvelun käytön yhteydessä) vai onko rekisterinpitäjä saanut ne jostain muusta lähteestä (esimerkiksi kolmannen osapuolen kautta tai haravoimalla). Molemmissa tapauksissa kohtuullisia odotuksia arvioitaessa olisi otettava huomioon toimenpiteet, joilla rekisteröidyille tiedotetaan käsittelytoimista.

95. Tekoälymallin käyttöönottovaiheessa on yhtä tärkeää ottaa huomioon rekisteröityjen kohtuulliset odotukset mallin erityisten ominaisuuksien puitteissa. Esimerkiksi sellaisten tekoälymallien osalta, jotka voivat mukautua annettujen syötteiden mukaan, voi olla merkityksellistä pohtia, ovatko rekisteröidyt olleet selvillä siitä, että he olivat antaneet henkilötietoja, jotta tekoälymalli voisi mukauttaa vastauksiaan heidän tarpeisiinsa ja he voisivat saada räätälöityjä palveluja. Lisäksi voi olla tärkeää harkita, vaikuttaako tämä käsittelytoimi ainoastaan rekisteröidyille tarjottavaan palveluun (esim. sisällön personointi tietylle käyttäjälle) vai käytetäänkö sitä kaikille asiakkaille tarjottavan palvelun muuttamiseen (esim. mallin yleiseen parantamiseen). Kuten kehittämisvaiheessa, voi myös olla erityisen tärkeää pohtia, onko rekisteröityjen ja rekisterinpitäjän välillä suora yhteys. Tällaisen suoran yhteyden avulla rekisterinpitäjä voi esimerkiksi antaa rekisteröidyille helposti tietoa käsittelytoiminnasta ja mallista, mikä voisi vaikuttaa näiden rekisteröityjen kohtuullisiin odotuksiin.

Lieventävät toimenpiteet

96. Jos rekisteröityjen edut, oikeudet ja vapaudet näyttävät olevan tärkeämpiä kuin rekisterinpitäjän tai kolmannen osapuolen tavoittelema(t) oikeutettu (oikeutetut) etu (edut), rekisterinpitäjä voi harkita lieventävien toimenpiteiden käyttöönottoa rajoittaakseen käsittelyn vaikutusta näihin rekisteröityihin. Lieventävät toimenpiteet ovat suojatoimia, jotka olisi räätälöitävä tapauksen olosuhteiden mukaan ja jotka riippuvat eri tekijöistä, kuten tekoälymallin aiotusta käyttötarkoituksesta. Tällaisten lieventävien toimenpiteiden olisi pyrittävä varmistamaan, että rekisterinpitäjän tai kolmannen osapuolen etuja ei syrjäytetä niin, että rekisterinpitäjä voisi vedota tähän oikeusperusteeseen.

97. Kuten Euroopan tietosuojaneuvoston oikeutettua etua koskevissa ohjeissa muistutetaan, lieventäviä toimenpiteitä ei pidä sekoittaa toimenpiteisiin, jotka rekisterinpitäjän on lain mukaan joka tapauksessa toteutettava yleisen tietosuoja-asetuksen noudattamisen varmistamiseksi katsomatta siihen, perustuuko käsittely yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohtaan[76]. Tämä on erityisen tärkeää toimenpiteissä, jotka edellyttävät esimerkiksi yleisen tietosuoja-asetuksen periaatteiden, kuten tietojen minimoinnin periaatteen, noudattamista.

98. Jäljempänä oleva toimenpideluettelo ei ole tyhjentävä eikä suuntaa-antava, ja toimenpiteiden täytäntöönpanoa olisi harkittava tapauskohtaisesti. Joitakin jäljempänä mainituista toimenpiteistä voidaan kuitenkin tilanteen mukaan edellyttää yleisen tietosuoja-asetuksen erityisten velvoitteiden noudattamiseksi. Jos näin ei ole, ne voidaan ottaa huomioon lisätakeina. Lisäksi jotkin jäljempänä mainituista toimenpiteistä liittyvät aloihin, joilla kehitys ja uudet kehityssuunnat ovat nopeita, mikä valvontaviranomaisten olisi otettava huomioon tiettyjen tapausten käsittelyssä.

99. **Tekoälymallien kehittämisvaiheessa** voidaan toteuttaa useita toimenpiteitä sekä ensimmäisen osapuolen että kolmannen osapuolen tietojen käsittelystä johtuvien riskien lieventämiseksi (myös verkkosivujen haravointikäytäntöihin liittyvien riskien vähentämiseksi). Edellä esitetyn perusteella eurooppalainen tietosuojalautakunta antaa joitakin esimerkkejä toimenpiteistä, joita voidaan

[76] Ohjeet 1/2024 henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, hyväksytty 8. lokakuuta 2024, 57 kohta.

toteuttaa tasapainotestissä havaittujen riskien lieventämiseksi ja jotka valvontaviranomaisten olisi otettava huomioon arvioidessaan tapauskohtaisesti tiettyjä tekoälymalleja.

100. **Tekniset toimenpiteet:**

a. Edellä 3.2.2 jaksossa mainitut toimenpiteet, jotka soveltuvat kyseessä olevien riskien lieventämiseen, jos kyseiset toimenpiteet eivät johda mallin anonymisointiin eikä niiden tarvitse noudattaa muita yleisen tietosuoja-asetuksen velvoitteita tai tarpeellisuustestiä (kohtuullisen edun arvioinnin toinen vaihe).

101. Näiden lisäksi muita asiaan liittyviä toimenpiteitä voivat olla muun muassa seuraavat:

b. Pseudonymisointitoimenpiteet: Tähän voisi sisältyä esimerkiksi toimenpiteitä, joilla estetään yksittäisiin tunnisteisiin perustuvien tietojen yhdistäminen. Nämä toimenpiteet eivät välttämättä ole asianmukaisia, jos valvontaviranomainen katsoo, että rekisterinpitäjä on osoittanut kohtuullisen tarpeen kerätä erilaisia tietoja tietystä henkilöstä kyseisen tekoälyjärjestelmän tai -mallin kehittämistä varten.

c. Toimenpiteet, joilla henkilötiedot peitetään tai korvataan keksityillä henkilötiedoilla koulutusaineistossa (esim. nimien ja sähköpostiosoitteiden korvaaminen keksityillä nimillä ja keksityillä sähköpostiosoitteilla). Toimenpide voi olla erityisen tarkoituksenmukainen silloin, kun tietojen olennainen sisältö ei ole merkityksellinen koko käsittelyn kannalta (esim. laajoihin kielimalleihin liittyvässä koulutuksessa).

102. **Toimenpiteet, jotka helpottavat yksilöiden oikeuksien käyttämistä:**

a. Kohtuullinen aika koulutustietoaineiston keräämisen ja käytön välillä. Tämä lisäsuojatoimi voi antaa rekisteröidyille mahdollisuuden käyttää oikeuksiaan kyseisen ajanjakson aikana. Kohtuullinen aika arvioidaan kunkin tapauksen olosuhteiden mukaan.

b. Ehdottoman poisjättäytymismahdollisuuden ehdottaminen alusta alkaen, esimerkiksi antamalla rekisteröidyille harkinnanvaraisen oikeuden vastustaa tietojen käsittelyä ennen käsittelyn aloittamista, jotta voidaan vahvistaa yksilöiden määräysvaltaa omiin tietoihinsa, mikä ylittää yleisen tietosuoja-asetuksen 21 artiklan edellytykset[77].

c. Annetaan rekisteröidyille mahdollisuus käyttää oikeuttaan tietojen poistamiseen, vaikka tietosuoja-asetuksen 17 artiklan 1 kohdassa lueteltuja erityisiä syitä ei voitaisi soveltaa[78].

d. Annetaan rekisteröidyille mahdollisuus toimittaa vaateita henkilötietojen regurgitaatiosta tai muistiin tallentamisesta sekä esittää olosuhteet ja tavat, joilla rekisterinpitäjät voivat jäljentää ja arvioida asiaankuuluvia vähemmän kehittyneitä tekniikoita vaateiden käsittelemiseksi.

103. **Läpinäkyvyyttä koskevat toimenpiteet:** joissakin tapauksissa lieventäviin toimenpiteisiin voi sisältyä toimenpiteitä, joilla lisätään tekoälymallin kehittämisen läpinäkyvyyttä. Yleisen tietosuoja-asetuksen velvoitteiden noudattamisen lisäksi jotkin toimenpiteet voivat auttaa poistamaan tietojen epäsymmetriaa ja parantaa rekisteröityjen ymmärrystä kehittämisvaiheessa tapahtuvasta käsittelystä:

a. Sellaisten julkisten ja helposti saatavilla olevien viestien julkaiseminen, joiden tiedot ovat kattavammat kuin yleisen tietosuoja-asetuksen 13 tai 14 artiklassa vaaditaan ja joissa esimerkiksi annetaan lisätietoja keräämisen kriteereistä ja kaikista käytetyistä

[77] Ks. edellinen alaviite.
[78] Ks. edellinen alaviite.

tietoaineistoista ottaen huomioon lasten ja haavoittuvassa asemassa olevien henkilöiden erityinen suojelu.

b. Vaihtoehtoiset tavat tiedottaa rekisteröidyille, kuten mediakampanjat, joissa rekisteröidyille tiedotetaan eri tiedotusvälineissä, sähköpostiin perustuva tiedotuskampanja, graafisen visualisoinnin käyttö, usein kysytyt kysymykset, avoimuusmerkinnät ja mallikortit, joiden systematisointi voisi jäsentää tekoälymalleja koskevien tietojen esitystapaa, sekä vapaaehtoiset vuotuiset avoimuusraportit.

104. **Erityiset lieventävät toimenpiteet verkkosivujen haravoinnin yhteydessä:** Kun otetaan huomioon, että kuten edellä mainittiin, verkkosivujen haravointiin liittyy erityisiä riskejä[79], tässä yhteydessä voitaisiin identifioida erityisiä lieventäviä toimenpiteitä. Valvontaviranomaiset voivat tarvittaessa ottaa ne huomioon edellä mainittujen lieventävien toimenpiteiden lisäksi tutkiessaan rekisterinpitäjiä, jotka harjoittavat verkkosivujen haravointia.

105. Erityistoimenpiteet, kun niitä ei tarvita oikeutetun edun arvioinnin toisessa vaiheessa, voivat osoittautua hyödyllisiksi riskin lieventämiseksi verkkosivujen haravoinnin yhteydessä. Näihin voivat kuulua **tekniset toimenpiteet**, kuten:

a. Lukuun ottamatta tietosisältö julkaisuista, joihin saattaa sisältyä henkilötietoja, joista voi aiheutua riskejä tietyille henkilöille tai henkilöryhmille (esim. henkilöille, jotka saattavat joutua väärinkäytön, vahingon tai jopa fyysisen haitan kohteeksi, jos tiedot julkistetaan).

b. Varmistetaan, että tiettyjä tietoluokkia ei kerätä tai että tietyt lähteet jätetään tiedonkeruun ulkopuolelle. Tämä voi tarkoittaa esimerkiksi tiettyjä verkkosivustoja, jotka ovat erityisen tunkeilevia niiden kohteen arkaluonteisuuden vuoksi.

c. Lukuun ottamatta keräämistä verkkosivustoilta (tai verkkosivustojen osioista), jotka ovat selvästi verkkosivujen haravoinnin kohteena, ja niiden sisällön uudelleenkäyttöä tekoälyä kouluttavien tietokantojen rakentamiseksi (esimerkiksi robotit.txt- tai ai.txt-tiedostot tai mitkä tahansa muut tunnustetut mekanismit, jotka ilmaisevat automaattisten hakurobottien käytön tai haravoinnin poissulkemisen).

d. Muiden asiaankuuluvien rajoitusten asettaminen keräämiselle, mahdollisesti myös määräaikoihin perustuvat kriteerit.

106. Esimerkit erityistoimenpiteistä, **joilla helpotetaan yksilöiden oikeuksien käyttöä ja avoimuutta** verkkoharavoinnin yhteydessä, voivat olla seuraavanlaisia: rekisterinpitäjän hallinnoiman poisjättäytymisluettelon (opt-out) luominen, joka mahdollistaa rekisteröidyille hänen tietojensa keräämisen vastustamisen tietyillä verkkosivustoilla tai verkkoalustoilla, jotka antavat tietoa, jotka tunnistavat heidät kyseisillä verkkosivustoilla, myös ennen tietojen keräämistä[80].

107. **Lieventämistoimenpiteitä koskevat erityiset näkökohdat käyttöönottovaiheessa:** Vaikka jotkin edellä mainituista toimenpiteistä voivat olla tärkeitä myös käyttöönottovaiheelle, Euroopan tietosuojaneuvosto esittää olosuhteista riippuen jäljempänä ei-tyhjentävän luettelon täydentävistä

[79] Nämä käytännöt voivat myös herättää lisäkysymyksiä, joita ei käsitellä tässä lausunnossa, ks. esimerkiksi Pagallo U., Ciani Sciolla J., *Anatomy of web data scraping: ethics, standards, and the troubles of the law.* European Journal of Privacy Law & Technologies, (2023) 2, s. 1–19, saatavilla osoitteessa https://doi.org/10.57230/EJPLT232PS.

[80] Ellei rekisterinpitäjä voi osoittaa, että käsittelyyn on olemassa huomattavan tärkeä ja perusteltu syy, joka syrjäyttää rekisteröidyn edut, oikeudet ja vapaudet, tai jos se on tarpeen oikeusvaateen laatimiseksi, esittämiseksi tai puolustamiseksi.

tukitoimenpiteistä, jotka voidaan toteuttaa ja jotka valvontaviranomaisten olisi arvioitava tapauskohtaisesti.

a. **Teknisiä toimenpiteitä** voidaan ottaa käyttöön esimerkiksi henkilötietojen tallentamisen, regurgitaation tai tuottamisen estämiseksi erityisesti generatiivisten tekoälymallien (kuten tuotossuodattimien) yhteydessä ja/tai laittoman uudelleenkäytön riskin vähentämiseksi yleiskäyttöisten tekoälymallien avulla (esim. tekoälyn tuottamien tuotosten digitaalinen vesileimaus).

b. **Toimenpiteet, joilla helpotetaan tai nopeutetaan yksilöiden oikeuksien käyttöä** käyttöönottovaiheessa enemmän kuin laissa edellytetään, ja jotka koskevat erityisesti, mutta eivät ainoastaan, oikeutta poistaa henkilötiedot mallin tuotostiedoista tai deduplikointia (kaksoiskappaleen poistamista) sekä koulutuksen jälkeisiä tekniikoita, joilla pyritään poistamaan tai häivyttämään henkilötietoja.

108. Tutkiessaan tietyn tekoälymallin käyttöönottoa valvontaviranomaisten olisi harkittava, onko rekisterinpitäjä julkaissut tekemänsä tasapainotestin, koska se voi edistää läpinäkyvyyttä ja kohtuullisuutta. Kuten Euroopan tietosuojaneuvoston oikeutettua etua koskevissa ohjeissa mainitaan, voidaan harkita muita toimenpiteitä, joilla rekisteröidyille voidaan antaa tietoa tasapainotestistä ennen henkilötietojen keräämistä[81]. Tietosuojaneuvosto toistaa myös[82], että huomioon otettava tekijä on se, onko rekisterinpitäjä ottanut tietosuojavastaavan mukaan prosessiin tarvittaessa.

3.4 Tekoälymallin kehittämisessä tapahtuneen lainvastaisen käsittelyn mahdollinen vaikutus tekoälymallin myöhemmän käsittelyn tai käytön lainmukaisuuteen

109. Tässä lausunnon jaksossa käsitellään pyynnön kysymystä 4. Kysymyksellä pyritään saamaan selvyyttä siihen, miten kehittämisvaiheessa tapahtuva lainvastainen käsittely mahdollisesti vaikuttaa myöhempään käsittelyyn (esimerkiksi tekoälymallin käyttöönottovaiheessa) tai mallin toimintaan. Kysymyksellä pyritään käsittelemään tilannetta, jossa tällainen tekoälymalli käsittelee mallissa säilytettäviä henkilötietoja (pyynnön kysymys 4 i), ja tilannetta, jossa tekoälymallin käyttöönottoon ei enää liity henkilötietojen käsittelyä (eli malli on anonyymi) (pyynnön kysymys 4 ii).

110. Ennen tiettyjen erityisskenaarioiden käsittelyä tietosuojaneuvosto esittää seuraavat yleiset näkökohdat.

111. Ensinnäkin tässä jaksossa esitetyissä selvennyksissä keskitytään henkilötietojen käsittelyyn kehittämisvaiheessa, jossa ei noudateta yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan a alakohdassa ja erityisesti yleisen tietosuoja-asetuksen 6 artiklassa vahvistettua lainmukaisuuden periaatetta (jäljempänä ’**lainvastaisuus**’)[83]. Vastaavasti tietosuojaneuvoston pohdinnoissa keskitytään siihen, miten käsittelyn lainvastaisuus kehitysvaiheessa vaikuttaa mallin myöhemmän käsittelyn tai toiminnan lainmukaisuuteen (eli yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan a alakohdan ja 6 artiklan noudattamiseen). Tietosuojaneuvosto huomauttaa kuitenkin, että kehittämisvaiheessa suoritettu käsittely voi johtaa myös muiden yleisen tietosuoja-asetuksen säännösten rikkomiseen, kuten

[81] Euroopan tietosuojaneuvoston henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 68 kohta.
[82] Euroopan tietosuojaneuvoston henkilötietojen käsittelystä yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan perusteella, versio 1.0, annettu 8. lokakuuta 2024, 12 kohta.
[83] Unionin tuomioistuimen tuomio 4 päivänä toukokuuta 2023 asiassa C-60/22, *Bundesrepublik Deutschland* (ECLI:EU:C:2023:373), 55–57 kohta.

avoimuuden puutteeseen rekisteröityjä kohtaan tai sisäänrakennettuun ja/tai oletusarvoiseen tietosuojaan, joita ei ole analysoitu tässä lausunnossa.

112. Toiseksi tätä kysymystä käsiteltäessä *keskeistä on osoitusvelvollisuusperiaate, jonka mukaan rekisterinpitäjien on vastattava* muun muassa[84] yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan ja yleisen tietosuoja-asetuksen 6 artiklan noudattamisesta ja osoitettava niiden noudattaminen. Tämä koskee myös tarvetta arvioida, mikä organisaatio on rekisterinpitäjä kyseessä olevassa käsittelytoimessa, ja sitä, esiintyykö yhteisrekisterinpitäjyyttä koskevia tilanteita (koska ne voivat liittyä erottamattomasti toisiinsa)[85]. Kun otetaan huomioon kunkin tapauksen tosiasiallisten olosuhteiden merkitys, mukaan lukien kunkin käsittelyyn osallistuvan osapuolen rooli, tietosuojaneuvoston näkemykset olisi ymmärrettävä yleisiksi huomautuksiksi, joita valvontaviranomaisten olisi arvioitava tapauskohtaisesti.

113. Kolmanneksi tietosuojaneuvosto korostaa, että tietosuoja-asetuksen 51 artiklan 1 kohdan mukaan valvontaviranomaiset ovat "*vastuussa [yleisen tietosuoja-asetuksen] soveltamisen valvonnasta luonnollisten henkilöiden perusoikeuksien ja -vapauksien suojaamiseksi käsittelyssä ja henkilötietojen vapaan liikkuvuuden helpottamiseksi unionissa*". Valvontaviranomaisilla on näin ollen toimivalta arvioida käsittelyn lainmukaisuutta ja käyttää yleisessä tietosuoja-asetuksessa annettuja valtuuksiaan kansallisen kehyksensä mukaisesti[86]. Tällaisissa tapauksissa valvontaviranomaisilla on harkintavalta arvioida mahdollinen rikkomus (mahdolliset rikkomukset) ja valita asianmukaiset, tarpeelliset ja oikeasuhteiset toimenpiteet yleisen tietosuoja-asetuksen 58 artiklassa mainituista toimenpiteistä ottaen huomioon kunkin yksittäisen tapauksen olosuhteet[87].

114. **Kun havaitaan rikkomus, valvontaviranomaiset voivat määrätä korjaavia toimenpiteitä, kuten määrätä rekisterinpitäjiä toteuttamaan kunkin tapauksen olosuhteet huomioon ottaen toimia alkuperäisen käsittelyn lainvastaisuuden korjaamiseksi.** Tällaisia keinoja voivat olla esimerkiksi sakon määrääminen, käsittelyn väliaikainen rajoittaminen, lainvastaisesti käsitellyn tietoaineiston osan poistaminen tai, jos tämä ei ole mahdollista, käsillä olevista tosiseikoista riippuen ja toimenpiteen oikeasuhteisuus huomioon ottaen, tekoälymallin kehittämisessä käytetyn koko tietoaineiston ja/tai itse tekoälymallin poistaminen. Arvioidessaan suunnitellun toimenpiteen oikeasuhteisuutta valvontaviranomaiset voivat ottaa huomioon toimenpiteet, joita rekisterinpitäjä voi soveltaa alkuperäisen käsittelyn lainvastaisuuden korjaamiseksi (esim. uudelleenkoulutus).

115. Tietosuojaneuvosto korostaa myös, että kun henkilötietoja käsitellään lainvastaisesti, rekisteröidyt voivat pyytää henkilötietojensa poistamista yleisen tietosuoja-asetuksen 17 artiklassa esitettyjen

[84] Unionin tuomioistuimen tuomio 4 päivänä toukokuuta 2023 asiassa C-60/22, *Bundesrepublik Deutschland* (ECLI:EU:C:2023:373), 53 kohta.

[85] Euroopan tietosuojaneuvoston ohjeet 7/2020 rekisterinpitäjän ja henkilötietojen käsittelijän käsitteistä yleisessä tietosuoja-asetuksessa, versio 2.1, 7. heinäkuuta 2021, 55 kohta.

[86] Erityiset kansalliset säännöt voidaan joutua ottamaan huomioon. Ks. esimerkiksi Italian tietosuojalain (196/2003) 2 pykälä, jossa säädetään, että tietosuojasääntöjen vastaisesti käsiteltyjä tietoja ei voida käyttää. Tämä ei rajoita muiden kansallisten oikeudellisten kehysten, kuten rikoslainsäädännön, soveltamista.

[87] Ks. tältä osin yleisen tietosuoja-asetuksen johdanto-osan 129 kappale sekä unionin tuomioistuimen tuomio 26 päivänä syyskuuta 2024, asiassa C-768–21, TR vastaan *Land Hessen* (ECLI:EU:C:2024:785), 37 kohta; unionin tuomioistuimen tuomio 7 päivänä joulukuuta 2023 yhdistetyissä asioissa C-26/22 ja C-64/22, SCHUFA Holding (Libération de reliquat de dette) (ECLI:EU:C:2023:958), 57 kohta; ja unionin tuomioistuimen tuomio 14 päivänä maaliskuuta 2024 asiassa C-46/23, *Újpesti Polgármesteri Hivatal* (ECLI:EU:C:2024:239), 34 kohta.

edellytysten mukaisesti ja että valvontaviranomaiset voivat määrätä poistamaan henkilötiedot *viran puolesta*[88].

116. Arvioidessaan, onko toimenpide asianmukainen, tarpeellinen ja oikeasuhteinen, valvontaviranomaiset voivat ottaa huomioon muun muassa rekisteröidyille aiheutuvat riskit, rikkomuksen vakavuuden, toimenpiteen teknisen ja taloudellisen toteutettavuuden sekä kyseessä olevien henkilötietojen määrän.

117. Lopuksi tietosuojaneuvosto muistuttaa, että valvontaviranomaisten yleisen tietosuoja-asetuksen nojalla toteuttamat toimenpiteet eivät rajoita toimivaltaisten viranomaisten tekoälysäädöksen ja/tai muiden sovellettavien oikeudellisten kehysten (esim. siviilioikeudellista vastuuta koskevan lainsäädännön) nojalla toteuttamia toimenpiteitä.

118. Seuraavissa jaksoissa tietosuojaneuvosto käsittelee kysymyksen 4 kolmea skenaariota. Niiden erot liittyvät siihen, säilytetäänkö mallin kehittämisessä käsiteltyjä henkilötietoja mallissa ja/tai suorittaako myöhemmän käsittelyn sama vai toinen rekisterinpitäjä.

3.4.1 Skenaario 1. Rekisterinpitäjä käsittelee henkilötietoja lainvastaisesti mallin kehittämiseksi, henkilötietoja säilytetään mallissa ja sama rekisterinpitäjä käsittelee niitä myöhemmin (esimerkiksi mallin käyttöönoton yhteydessä).

119. Tämä skenaario liittyy pyynnössä kysymyksen 4 i alakohtaan tilanteessa, jossa rekisterinpitäjä käsittelee henkilötietoja lainvastaisesti (eli ei noudata yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan a alakohtaa eikä 6 artiklaa) kehittääkseen tekoälymallia, ja tekoälymallissa säilytetään tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön liittyviä tietoja, eikä se näin ollen ole anonyymi. Tämän jälkeen sama rekisterinpitäjä käsittelee henkilötietoja (esimerkiksi mallin käyttöönoton yhteydessä). Tämän skenaarion osalta Euroopan tietosuojaneuvosto esittää seuraavat näkökohdat.

120. Valvontaviranomaisen valtuudet määrätä alkuperäiseen käsittelyyn liittyviä korjaavia toimenpiteitä (kuten edellä 113, 114 ja 115 kohdassa on selitetty) vaikuttaisi periaatteessa myöhempään käsittelyyn (esim. jos valvontaviranomainen määrää rekisterinpitäjän poistamaan laittomasti käsitellyt henkilötiedot, nämä korjaavat toimenpiteet eivät antaisi viimeksi mainitulle mahdollisuutta käsitellä myöhemmin toimenpiteiden kohteena olevia henkilötietoja).

121. Mitä tulee erityisesti kehittämisvaiheen laittoman käsittelyn vaikutukseen myöhemmässä käsittelyssä (esimerkiksi käyttöönottovaiheessa), Euroopan tietosuojaneuvosto muistuttaa, että valvontaviranomaisten tehtävänä on tehdä tapauskohtainen analyysi, jossa otetaan huomioon kunkin tapauksen erityisolosuhteet.

122. **Se, liittyykö kehitys- ja käyttöönottovaiheisiin erillisiä tarkoituksia (ja ovatko ne siten erillisiä käsittelytoimia) ja missä määrin oikeusperusteen puuttuminen alkuperäisestä käsittelytoimesta vaikuttaa myöhemmän käsittelyn laillisuuteen, olisi arvioitava tapauskohtaisesti tapauksen asiayhteydestä riippuen.**

123. Esimerkiksi yleisen tietosuoja-asetuksen 6 artiklan 1 kohdan f alakohdan oikeusperusteeseen liittyen, kun myöhempi käsittely perustuu oikeutettuun etuun, oikeutetun edun arvioinnissa olisi otettava

[88] Tältä osin EDPB:n lausunto 39/2021 siitä, voisiko tietosuoja-asetuksen 58 artiklan 2 kohdan g alakohta toimia oikeusperusteena, jonka nojalla valvontaviranomainen voi määrätä viran puolesta henkilötietojen poistamisesta tilanteessa, jossa rekisteröity ei ole esittänyt tällaista pyyntöä, 28 kohta. Ks. tältä osin myös unionin tuomioistuimen tuomio 14 päivänä maaliskuuta 2024, asiassa C-46/23, Újpesti *Polgármesteri Hivatal* (ECLI:EU:C:2024:239), 42 kohta.

huomioon se, että alkuperäinen käsittely on ollut lainvastaista (esim. rekisteröidyille aiheutuvien riskien osalta tai sen osalta, että rekisteröidyt eivät ehkä odota tällaista myöhempää käsittelyä). Näissä tapauksissa käsittelyn lainvastaisuus kehittämisvaiheessa voi vaikuttaa myöhemmän käsittelyn lainmukaisuuteen.

3.4.2 Skenaario 2. Rekisterinpitäjä käsittelee henkilötietoja lainvastaisesti mallin kehittämiseksi, henkilötietoja säilytetään mallissa ja toinen rekisterinpitäjä käsittelee niitä mallin käyttöönoton yhteydessä.

124. Skenaario liittyy pyynnön kysymykseen 4 i. Se eroaa skenaariosta 1 (tämän lausunnon jakso 3.4.1), koska toinen rekisterinpitäjä käsittelee henkilötietoja myöhemmin tekoälymallin käyttöönoton yhteydessä.

125. Tietosuojaneuvosto muistuttaa, että näille eri toimijoille tietosuojakehyksen puitteissa annettujen roolien varmistaminen on tärkeä vaihe sen määrittämiseksi, mitä yleisen tietosuoja-asetuksen mukaisia velvoitteita sovelletaan ja kuka on vastuussa näistä velvoitteista, ja että yhteisrekisterinpidon tilanteet olisi myös otettava huomioon arvioitaessa kunkin osapuolen vastuita tietosuoja-asetuksen nojalla. Siksi jäljempänä esitetyt huomautukset on katsottava yleisiksi tekijöiksi, jotka valvontaviranomaisten on tarvittaessa otettava huomioon. Skenaarion 2 osalta tietosuojaneuvosto esittää seuraavat näkökohdat.

126. Ensinnäkin on muistettava, että yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan a alakohdan mukaan, luettuna yhdessä yleisen tietosuoja-asetuksen 5 artiklan 2 kohdan kanssa, jokaisen rekisterinpitäjän on varmistettava suorittamansa käsittelyn lainmukaisuus ja pystyttävä osoittamaan se. Siksi valvontaviranomaisten olisi arvioitava i) tekoälymallin alun perin kehittäneen rekisterinpitäjän ja ii) tekoälymallin hankkineen ja henkilötietoja itse käsittelevän rekisterinpitäjän suorittaman käsittelyn lainmukaisuutta.

127. Toiseksi edellä 113, 114 ja 115 kohdassa esitetyt näkökohdat ovat tässä tapauksessa merkittäviä, kun on kyse valvontaviranomaisten valtuuksista puuttua alkuperäiseen käsittelyyn. Yleisen tietosuoja-asetuksen 17 artiklan 1 kohdan d alakohta (lainvastaisesti käsiteltyjen tietojen poistaminen) ja 19 artikla (henkilötietojen oikaisua tai poistoa tai käsittelyn rajoitusta koskeva ilmoitusvelvollisuus) voivat tapauksen olosuhteista riippuen olla merkityksellisiä tässä yhteydessä, esimerkiksi kun on kyse ilmoituksesta, joka mallia kehittävän rekisterinpitäjän pitäisi antaa mallin käyttöön ottavalle rekisterinpitäjälle.

128. Kolmanneksi valvontaviranomaisten olisi arvioitava tapauskohtaisesti alkuperäisen käsittelyn lainvastaisuuden mahdollinen vaikutus toisen rekisterinpitäjän suorittamaan myöhempään käsittelyyn.

129. **Valvontaviranomaisten pitäisi ottaa huomioon, onko mallin käyttöön ottava rekisterinpitäjä tehnyt asianmukaisen arvioinnin osana osoitusvelvollisuuttaan [89] osoittaakseen yleisen tietosuoja-asetuksen 5 artiklan 1 kohdan a alakohdan ja 6 artiklan noudattamisen, jotta voidaan varmistaa, että tekoälymallia ei ole kehitetty käsittelemällä henkilötietoja lainvastaisesti.** Valvontaviranomaisten olisi otettava arvioinnissa huomioon, onko rekisterinpitäjä arvioinut joitakin ei-tyhjentäviä kriteerejä, kuten tietolähteen ja sen, onko tekoälymallin taustalla yleisen tietosuoja-asetuksen rikkominen, erityisesti, jos valvontaviranomainen tai tuomioistuin on tähän päätynyt, jotta

[89] Yleisen tietosuoja-asetuksen 5 artiklan 2 kohta ja 24 artikla.

mallia käyttävä rekisterinpitäjä ei voisi jättää huomiotta sitä, että alkuperäinen käsittely on ollut lainvastaista.

130. Rekisterinpitäjän olisi esimerkiksi harkittava, onko tiedot saatu henkilötietojen tietoturvaloukkauksen yhteydessä tai onko valvontaviranomainen tai tuomioistuin havainnut käsittelyssä rikkomisen. **Rekisterinpitäjän arvioinnin tarkkuus ja valvontaviranomaisten odottama yksityiskohtaisuus voivat vaihdella eri tekijöiden mukaan, mukaan lukien niiden riskien tyyppi ja laajuus, joita tekoälymallissa tapahtuva käsittely sen käyttöönoton aikana aiheuttaa suhteessa rekisteröityihin, joiden tietoja on käytetty mallin kehittämiseen.**

131. Tietosuojaneuvosto toteaa, että tekoälysäädös edellyttää, että korkean riskin tekoälyjärjestelmien tarjoajien on laadittava EU:n vaatimustenmukaisuusvakuutus[90], ja että tällainen vakuutus sisältää lausuman siitä, että kyseinen tekoälyjärjestelmä on EU:n tietosuojalainsäädännön mukainen[91]. Tietosuojaneuvosto toteaa, että tällainen oma ilmoitus ei välttämättä ole ratkaiseva toteamus yleisen tietosuoja-asetuksen mukaisesta säännösten noudattamisesta. Valvontaviranomaiset voivat kuitenkin ottaa sen huomioon tutkiessaan tiettyä tekoälymallia.

132. Samat näkökohdan kuin edellä 123 kohdassa ovat merkityksellisiä myös nyt käsiteltävässä asiassa. Kun valvontaviranomaiset tarkistavat, onko rekisterinpitäjä arvioinut oikeutetun edun asianmukaisuuden suorittamansa käsittelyn oikeusperusteena ja miten se on arvioitu, alkuperäisen käsittelyn lainvastaisuus olisi otettava huomioon osana oikeutetun edun arviointia esimerkiksi arvioimalla mahdollisia riskejä, joita voi aiheutua niille rekisteröidyille, joiden henkilötietoja on käsitelty lainvastaisesti mallin kehittämistarkoituksessa. Tasapainotestissä on otettava asianmukaisesti huomioon erilaiset tekijät, jotka ovat luonteeltaan joko teknisiä (esim. mallin kehittämisen aikaisten suodattimien tai pääsyä koskevat rajoitukset, joita seuraava rekisterinpitäjä ei voi kiertää tai joihin tämä ei voi vaikuttaa ja jotka saattavat estää henkilötietoihin pääsyn tai niiden luovuttamisen) tai oikeudellisia (esim. alkuperäisen käsittelyn lainvastaisuuden luonne ja vakavuus).

3.4.3 Skenaario 3. Rekisterinpitäjä käsittelee henkilötietoja mallin kehittämiseksi lainvastaisesti ja varmistaa sitten, että malli anonymisoidaan, ennen kuin sama tai toinen rekisterinpitäjä aloittaa uuden henkilötietojen käsittelyn käyttöönoton yhteydessä.

133. Tämä skenaario liittyy pyynnön 4 kohdan ii alakohtaan ja viittaa tapaukseen, jossa rekisterinpitäjä käsittelee henkilötietoja lainvastaisesti tekoälymallin kehittämiseksi, mutta tekee sen tavalla, jolla varmistetaan, että henkilötiedot anonymisoidaan, ennen kuin sama tai toinen rekisterinpitäjä aloittaa toisen henkilötietojen käsittelyn käyttöönoton yhteydessä. Ensinnäkin tietosuojaneuvosto muistuttaa, että valvontaviranomaiset ovat toimivaltaisia ja niillä on valtuudet puuttua asiaan mallin anonymisointiin liittyvän käsittelyn sekä käsittelyn kehittämisvaiheen aikana. Näin ollen valvontaviranomaiset voivat tapauksen erityisolosuhteista riippuen määrätä korjaavia toimenpiteitä tämän alustavan käsittelyn osalta (kuten edellä 113, 114 ja 115 kohdassa on selitetty).

134. Jos voidaan osoittaa, että tekoälymallin myöhempi toiminta ei sisällä henkilötietojen käsittelyä, tietosuojaneuvosto katsoo, että yleistä tietosuoja-asetusta ei sovelleta[92]. Näin ollen alkuperäisen käsittelyn lainvastaisuus ei saisi vaikuttaa mallin myöhempään toimintaan. Tietosuojaneuvosto korostaa, että pelkästään väite mallin anonymiteetista ei riitä vapauttamaan sitä yleisen tietosuoja-asetuksen soveltamisesta, ja toteaa, että valvontaviranomaisten olisi arvioitava mallia

[90] Tekoälysäädöksen 16 artiklan g alakohta ja 47 artikla.
[91] Tekoälysäädöksen liitteen V 5 kohta.
[92] Yleisen tietosuoja-asetuksen johdanto-osan 26 kappale.

tapauskohtaisesti ottaen huomioon tietosuojaneuvoston näkökohdat pyynnön ensimmäiseen kysymykseen vastaamiseksi.

135. **Kun rekisterinpitäjät käsittelevät myöhemmin käyttöönottovaiheen aikana kerättyjä henkilötietoja sen jälkeen, kun malli on anonymisoitu, yleistä tietosuoja-asetusta sovelletaan näihin käsittelytoimiin. Mitä tulee yleiseen tietosuoja-asetukseen, alkuperäisen käsittelyn lainvastaisuus näissä tapauksissa ei saisi vaikuttaa käyttöönottovaiheessa suoritetun käsittelyn lainmukaisuuteen.**

4 Loppuhuomautukset

136. Tämä lausunto osoitetaan kaikille valvontaviranomaiselle ja julkaistaan yleisen tietosuoja-asetuksen 64 artiklan 5 kohdan b alakohdan mukaisesti.

Euroopan tietosuojaneuvoston puolesta

Puheenjohtaja

Anu Talus