EU:n tekoälyasetus – tekoälyjärjestelmien ja yleiskäyttöisten tekoälymallien turvallisuussäädös*

Teknologiateollisuus ry

12.2.2024

**Tässä diasetissä esitetyt säännöstulkinnat ovat alustavia ja niihin voi tulla muutoksia lopullisen asetustekstin tultua julki ja tulkintaa tukevan ymmärryksen lisäännyttyä.*

Sisällys

Tiivistelmä

1. Soveltamisala

- Asetus koskee EU:ssa markkinoille saatettavia ja EU:ssa käyttöönotettavia tekoälyjärjestelmiä sekä yleiskäyttöisiä tekoälyämalleja, kuten suuria kielimalleja, joille monet tekoälyjärjestelmät pohjautuvat.
- Yksinkertaiset ohjelmistot ja ihmisen asettamia sääntöjä suoraviivaisesti seuraavat päätöksentekojärjestelmät eivät tule asetuksen piiriin.
- Ilmaiset, avoimen lähdekoodin järjestelmät ja mallit tulevat pääosin säädöksen piiriin.
- Asetuksen ulkopuolelle jäävät tutkimus- ja kehitysvaiheessa olevat järjestelmät ja mallit sekä järjestelmät, jotka on tarkoitettu yksinomaan sotilaalliseen, puolustukselliseen tai kansallisen turvallisuuden käyttöön.

2. Turvallisuussäädös

- Asetuksen tavoitteena on suojata terveyttä, turvallisuutta ja perusoikeuksia tekoälyjärjestelmien ja yleiskäyttöisten tekoälymallien käytöstä aiheutuvilta riskeiltä ja haitoilta.
- Säädös pyrkii asettamaan tekoälyjärjestelmien ja -mallien vaatimukset riskiperusteisesti. Olennaista on tarjoajana tai käyttöönottajana tietää, mihin riskiluokkaan järjestelmä tai malli kuuluu, jotta voi täyttää velvoitteensa.

3. Tekoälyjärjestelmien riskiluokittelu

- Kestämättömiä riskejä sisältävien järjestelmien käyttö kielletään. Esimerkkinä voidaan mainita sosiaalinen pisteytys. Joidenkin kiellettyjen järjestelmien käyttö on sallittua tietyissä tarkkaan rajatuissa tapauksissa.
- Korkean riskin järjestelmien käyttö on sallittua, kunhan järjestelmä täyttää sille säädetyt vaatimukset. Korkean riskin piiri kattaa tuotteita ja niiden turvakomponentteja, jotka ovat EU:n tuoteturvallisuuslainsäädännön piirissä, kuten koneet. Lisäksi korkean riskin piiriin luetaan muita määriteltyjä käyttötapauksia, kuten rekrytointi ja kriittisen digitaalisen infrastruktuurin hallinnointi. Korkean riskin vaatimuksista on mahdollisuus poiketa tietyissä määritellyissä tapauksissa.
- Rajoittuneita riskejä käsittävien generatiivisten ja muiden järjestelmien tarjoamiselle ja käytölle asetetaan tiettyjä avoimuusvaatimuksia. Ihmisen on esimerkiksi saatava tieto siitä, että hän kommunikoi tekoälyjärjestelmän kanssa.
- Vähäisen riskin järjestelmille ei aseteta vaatimuksia. Suurin osa järjestelmistä, mukaan lukien valtaosa teollisuuden käyttötapauksista, lukeutuu tähän luokkaan.

4. Yleiskäyttöisten tekoälymallien riskiluokittelu

- Kaikkia malleja koskevat tietyt avoimuus- ja informointivaatimukset. Näitä ovat muun muassa mallin tekoälyjärjestelmäänsä integroivien jatkokehittäjien informoiminen mallin ominaisuuksista sekä riittävän yksityiskohtaisen tiivistelmän laatiminen mallin koulutukseen käytetyistä sisällöistä.
- Järjestelmäriskejä sisältävien mallien, jotka ovat tyypillisesti isoimpia perustamalleja, on lisäksi täytettävä tietyt lisävaatimukset, mukaan lukien mallin arviointi ja vastatestaus.

5. Asteittainen soveltamisaikataulu

- Kielletyt käyttötapaukset soveltuvat, kun asetuksen voimaan astumisesta on kulunut 6 kuukautta.
- Yleiskäyttöisten tekoälymällien vaatimuksia ryhdytään soveltamaan, kun voimaan astumisesta on kulunut 12 kuukautta.
- Rajoittuneen riskin järjestelmien ja tiettyjen korkean riskin järjestelmien vaatimuksia aletaan soveltamaan, kun voimaan astumisesta on kulunut 24 kuukautta.
- Loput korkean riskin järjestelmät tulevat sovellettaviksi 36 kuukauden kuluttua säädöksen voimaan astumisesta.

Soveltamisala ja määritelmät

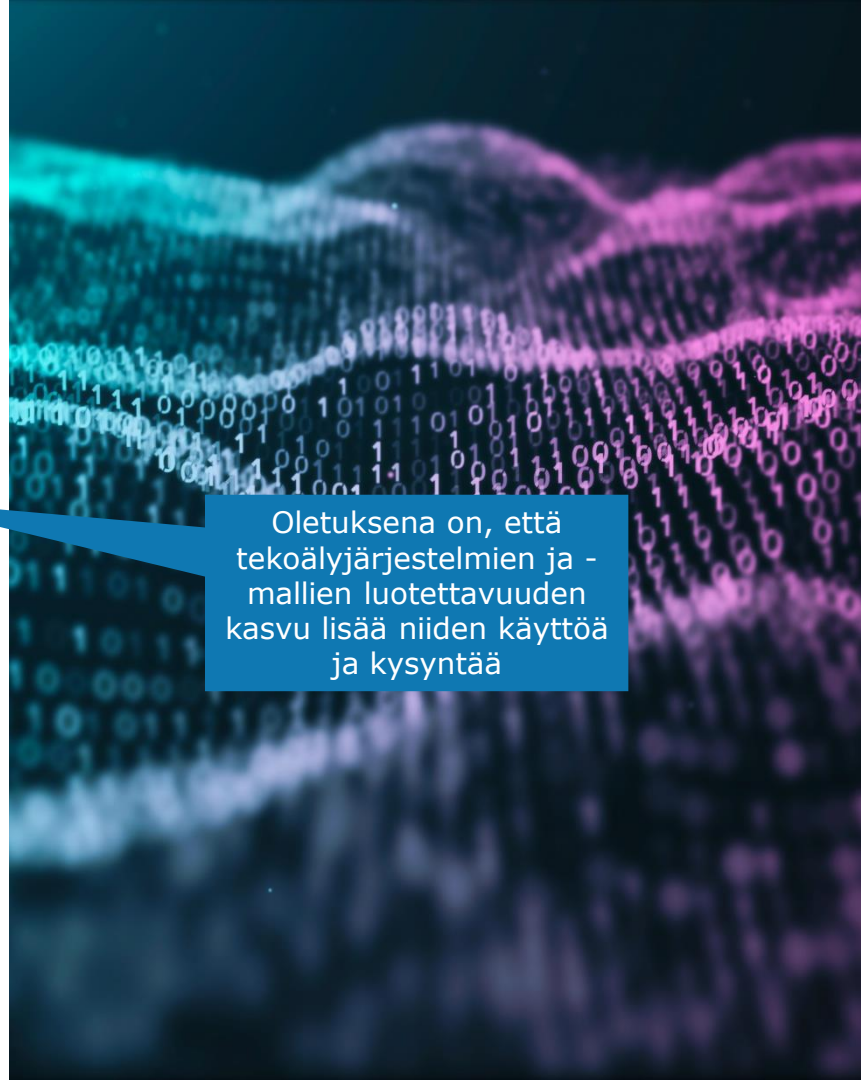
Mikä ihmeen tekoälyasetus?

Tekoälyjärjestelmien ja yleiskäyttöisten tekoälymallien turvallisuussäädös

- Tavoitteena on suojata terveyttä, turvallisuutta ja perusoikeuksia tekoälyjärjestelmien ja yleiskäyttöisten tekoälymallien käytöstä aiheutuvilta riskeiltä ja haitoilta.

- Sovelletaan sellaisenaan EU:ssa kaikilla sektoreilla ja toimialoilla.

- Säädös pyrkii asettamaan tekoälyjärjestelmien ja -mallien vaatimukset riskiperusteisesti.

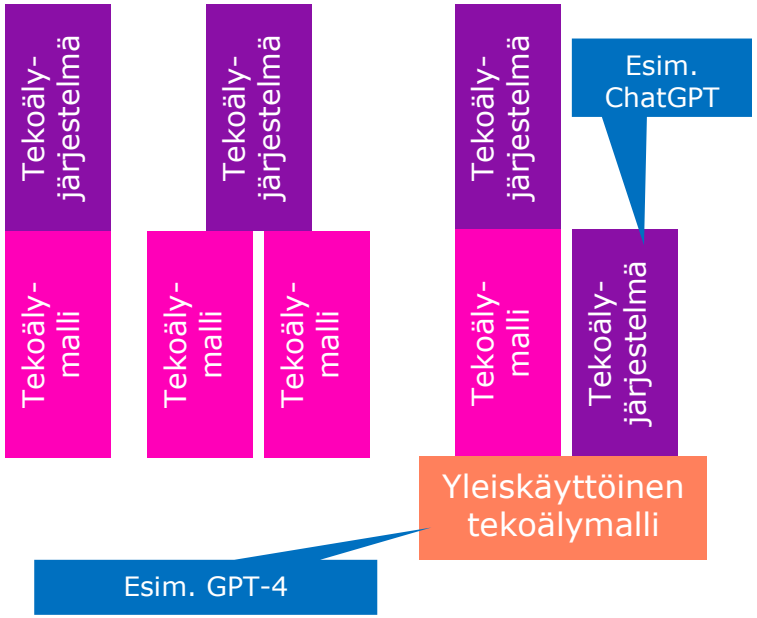


Asetuksen viitekehys:

Tekoälyjärjestelmät ja tekoälymallit

- Tekoälyjärjestelmä on sovellus, jolla on tietty käyttötarkoitus.
  - Järjestelmä voi rakentua yhden tai useamman yleiskäyttöisen tai kapean tekoälymallin varaan.

- Tekoälymalli on tekoälyjärjestelmän moottori, joka määrittää, miten järjestelmä käsittelee tietoa tietyissä tehtävissä, kuten kuvan tunnistuksessa tai kielen kääntämisessä.
  - On kapean tehtäväkentän omaavia malleja ja yleiskäyttöisiä malleja, kuten suuret kielimallit, joiden tehtäväkenttä on laaja.
  - Yleiskäyttöistä mallia voidaan jatkokehittää rajatummaksi malliksi.



Määritelmä:

Tekoälyjärjestelmä

- Suunniteltu toimimaan vaihtelevalla autonomian tasolla.
- Voi kyetä mukautumaan käyttöönoton jälkeen.
- Tavoitteitaan varten päättelee vastaanottamastaan syötteestä, kuinka luodaan tuotoksia, kuten ennusteita, sisältöä, suosituksia tai päätöksiä.
- Tuotokset vaikuttavat fyysiseen tai virtuaaliseen ympäristöön.

An AI system *is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.*

Määritelmä ei kata:
- Perinteiset yksinkertaiset ohjelmistojärjestelmät
- Automaattiset päätöksentekojärjestelmät, jotka noudattavat ihmisen asettamia sääntöjä

Määritelmä:

Yleiskäyttöinen tekoälymalli

Esim. suuret kielimallit ja multimodaaliset perustamallit

- Tyypillisesti koulutettu suurella määrällä dataa ja käsittää suuren määrän parametrejä.
- Kykenee suorittamaan pätevästi laajan määrän erilaisia tehtäviä.
- Voidaan integroida erilaisiin järjestelmiin tai sovelluksiin.

General-purpose AI model *is an AI model, including when trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable to competently perform a wide range of distinct tasks regardless of the way the model is released on the market and that can be integrated into a variety of downstream systems or applications.*

Voidaan saattaa markkinoille esimerkiksi:
- Kirjaston tai ohjelmointirajapinnan kautta
- Latauksena
- Fyysisenä kopiona

Keskeiset toimijat

Tarjoaja (Provider)

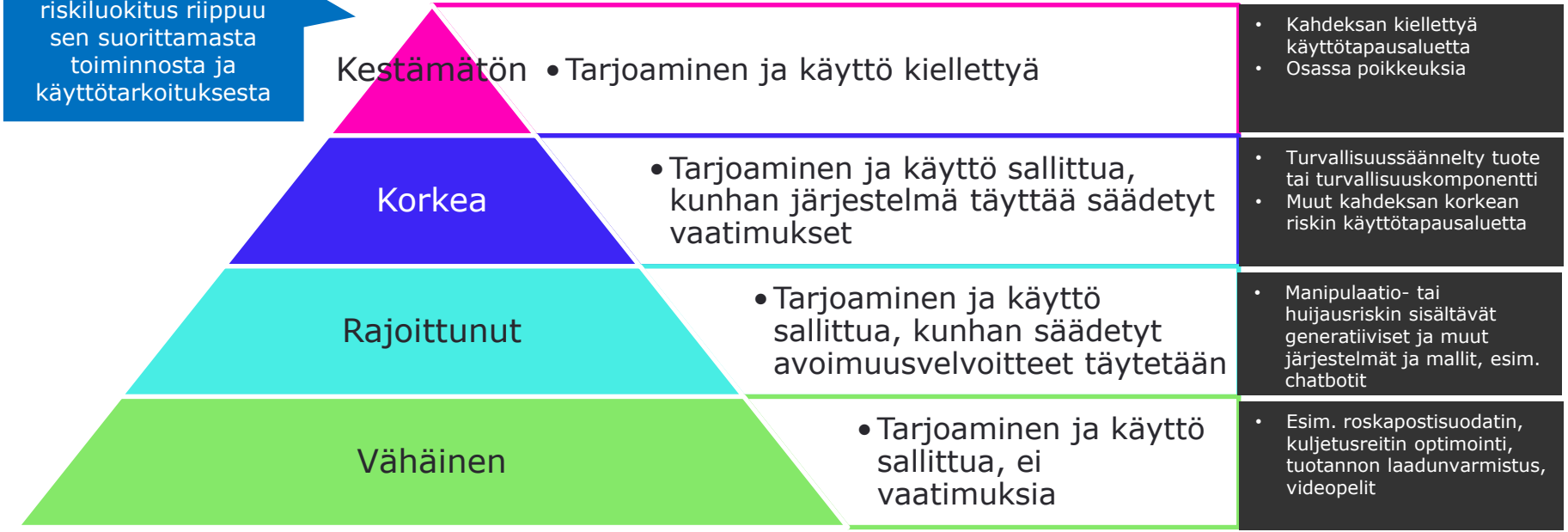
- Kehittää tai kehityttää tekoälyjärjestelmän tai yleiskäyttöisen tekoälymallin ja saattaa sen markkinoille tai ottaa sen käyttöön omalla nimellä tai tavaramerkillä joko maksua vastaan tai ilmaiseksi
- Riippumaton sijoittautumispaikasta
- Yritys tai muu oikeushenkilö, julkinen toimija, luonnollinen henkilö

Käyttöönottaja (Deployer)

- Käyttää tekoälyjärjestelmää
- Sijoittautunut EU-alueelle tai käytetyn järjestelmän tuotos kohdistuu EU-alueelle
- Yritys tai muu oikeushenkilö, julkinen toimija, luonnollinen henkilö
- Luonnollisen henkilön yksityinen, ei-ammattimainen käyttö ei lukeudu soveltamisalaan

Riskiluokittelu

Tekoälyjärjestelmän riskiluokitus riippuu sen suorittamasta toiminnosta ja käyttötarkoituksesta



Yleiskäyttöisille tekoälymalleille kaksitasoinen riskiluokittelu

Järjestelmäriskejä sisältävät mallit	• Lisävaatimukset
Kaikki yleiskäyttöiset mallit	• Avoimuus- ja informointivaatimukset

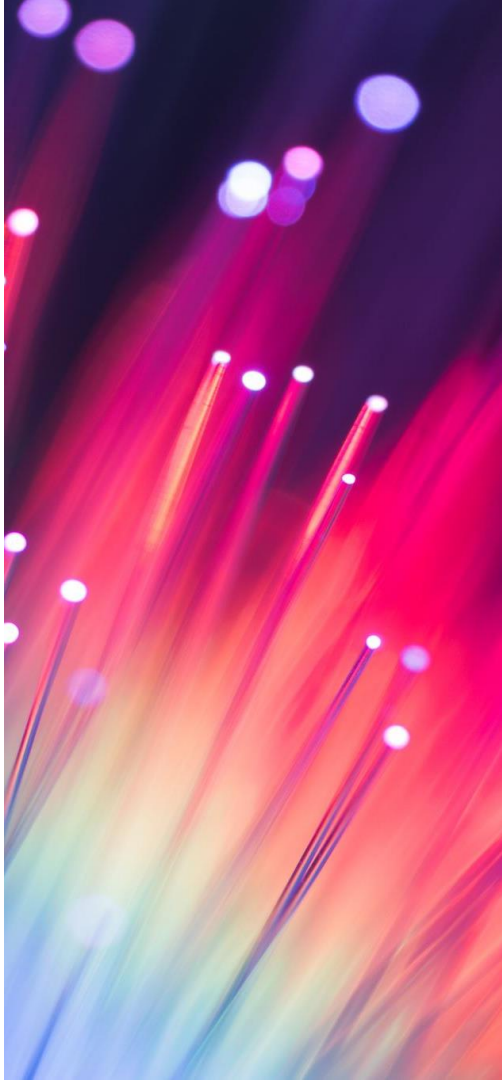
Asetuksen piiriin eivät tule

- Tieteelliseen tutkimus- ja kehityskäyttöön tarkoitetut tekoälyjärjestelmät ja -mallit.

- Tutkimus- ja kehitysvaiheessa olevat järjestelmät ja mallit, joita ei ole vielä saatettu markkinoille tai otettu käyttöön.
  - Tämä poikkeus ei koske järjestelmän tai mallin testausta todellisissa olosuhteissa.

- Yksinomaan sotilaalliseen, puolustukselliseen tai kansallisen turvallisuuden käyttöön tarkoitetut järjestelmät.

- Luonnolliset henkilöt, jotka käyttävät järjestelmää yksityisessä, ei-ammattimaisessa tarkoituksessa.



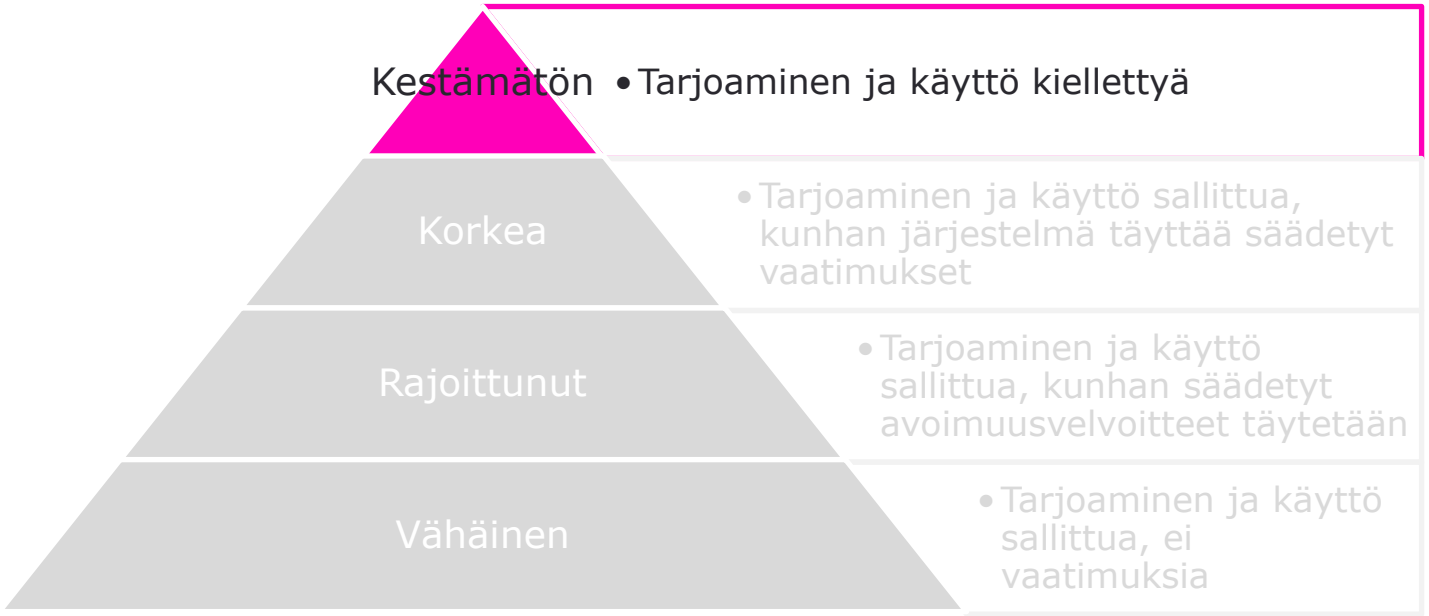
Avoimen lähdekoodin tekoälyjärjestelmät

- Jos ilmaisilla ja avoimen lähdekoodin lisensseillä tarjotut järjestelmät ovat korkean riskin järjestelmiä, on niiden noudatettava asetuksen vaatimuksia.

- Kielletyt käyttötapaukset koskevat myös avoimen lähdekoodin järjestelmiä.

- Ihmisen kanssa vuorovaikuttavien avoimen lähdekoodin järjestelmien on noudatettava avoimuusvaatimuksia.

Kielletyt käyttötapaukset



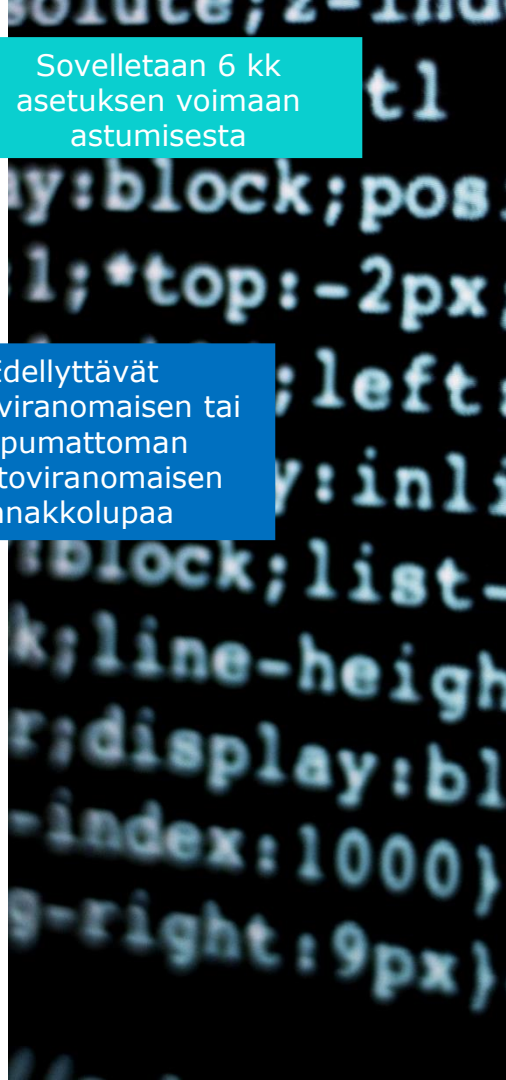
Järjestelmäriskejä sisältävät mallit	• Lisävaatimukset
Kaikki yleiskäyttöiset mallit	• Avoimuus- ja informointivaatimukset

Kielletyt käyttötapaukset 1/2

- Reaaliaikainen biometrinen etätunnistaminen julkisissa tiloissa lainvalvontatarkoituksiin
  - Ei kata:
    - Terrori-iskun ehkäisy tai siihen vastaaminen
    - Kadonneiden tai rikoksen uhrien etsintä
    - Tiettyihin vakaviin rikoksiin liittyvä lainvalvonta
- Haitalliset alitajuiset tekniikat
- Haavoittuvien ryhmien hyväksikäyttö
  - Ikä, vamma, sosiaalinen tai taloudellinen tilanne
- Sosiaalinen pisteytys, joka johtaa henkilöiden tai ryhmien haitalliseen tai epäsuotuisaan kohteluun

Sovelletaan 6 kk asetuksen voimaan astumisesta

Edellyttävät oikeusviranomaisen tai riippumattoman hallintoviranomaisen ennakkolupaa



Kielletyt käyttötapaukset 2/2

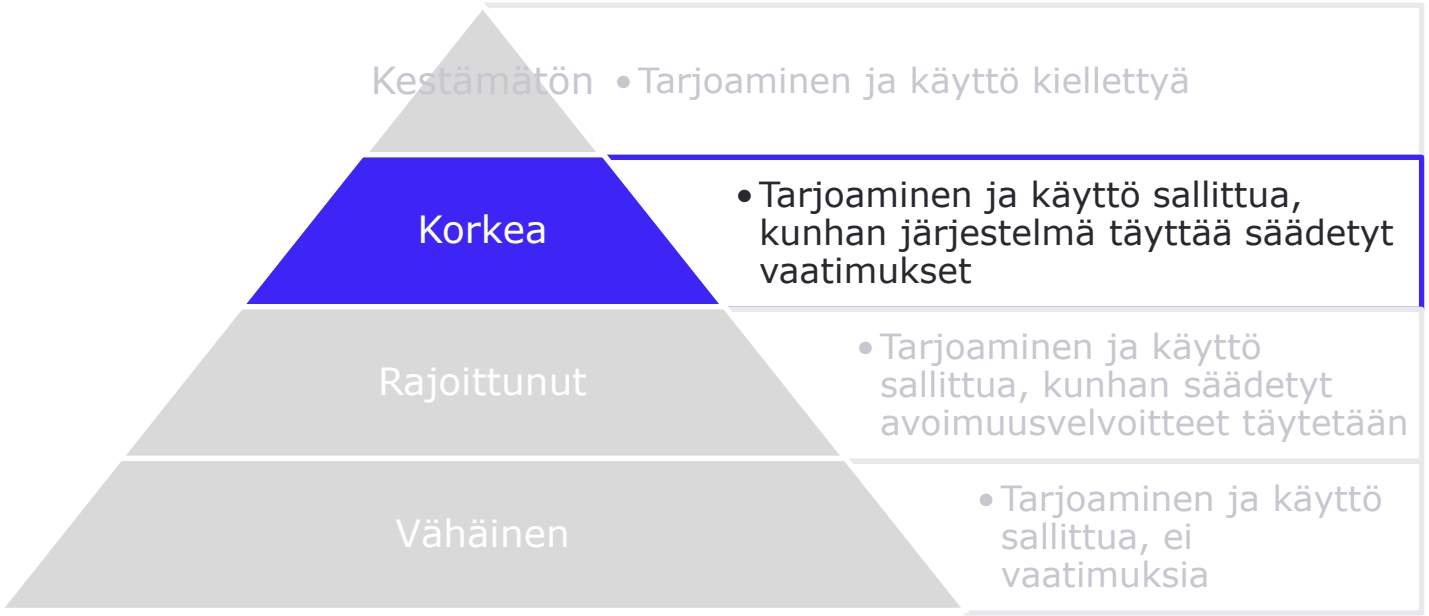
- Henkilön luokittelu tämän biometristen tietojen perusteella rodun, poliittisten mielipiteiden, ammattiliittojen jäsenyyden, uskonnollisen tai filosofisen vakaumuksen tai seksuaalisen suuntautumisen päättelemiseksi
  - Ei kata laillisesti hankittujen biometristen tietojoukkojen merkintöjä tai suodattamista esimerkiksi lainvalvontatarkoituksissa.

- Profilointiin perustuva rikosriskin ennakointi tai arviointi

- Tunteiden tunnistaminen työpaikalla tai oppilaitoksessa
  - Ei koske lääketieteellisiä tai turvallisuuteen liittyviä käyttötarkoituksia

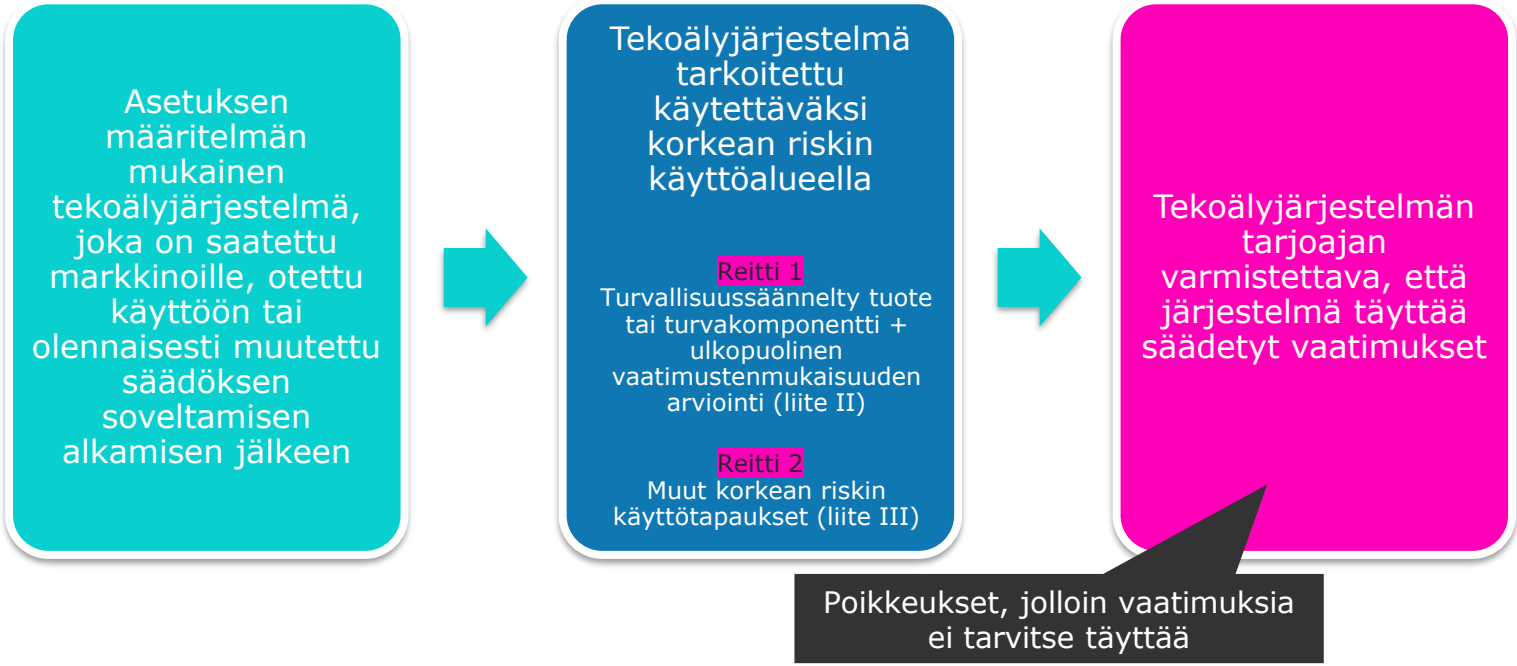
- Kasvojentunnistustietokantojen luominen tai laajentaminen kasvokuvien kohdistamattoman kaapimisen avulla internetistä tai valvontakamera-aineistosta

Korkean riskin käyttötapaukset ja vaatimukset



Järjestelmäriskejä sisältävät mallit	• Lisävaatimukset
Kaikki yleiskäyttöiset mallit	• Avoimuus- ja informointivaatimukset

Reitit korkean riskin vaatimusten piiriin



Reitti 1

Sovelletaan 36 kk asetuksen voimaan astumisesta

Korkean riskin käyttötapaukset:

Turvallisuussäännellyt tuotteet ja turvakomponentit

- Tekoälyjärjestelmä, joka

1. on tuote tai tuotteen turvakomponentti, joka kuuluu EU:n tuoteturvallisuuslainsäädännön piiriin

2. JA jolle on tuon sääntelyn perusteella suoritettava kolmannen osapuolen vaatimustenmukaisuuden arviointi.

Turvallisuussäädökset luetellaan asetuksen liitteessä II:
NLF-mukautetut säädökset (osio A)
+ muut harmonisoidut säädökset (osio B)

Reitti 1

Tuoteturvasääntelyn kautta tekoälyasetuksen piiriin tulevia tuoteryhmiä

- Koneet
- Hissit
- Lääkinnälliset laitteet
- Radiolaitteet
- Painelaitteet
- Veneet ja vesiskootterit
- Henkilösuojaimet
- Kaasumaisia polttoaineita polttavat laitteet
- Lelut
- Siviililentokoneet
- Kaksi-, kolmi- ja nelipyöräiset ajoneuvot
- Maa- ja metsätalousajoneuvot
- Laivavarusteet
- Moottoriajoneuvot ja niiden perävaunut

Turvallisuus-säädökset luetellaan asetuksen liitteessä II

Reitti 2

Muut korkean riskin käyttötapaukset 1/3

Sovelletaan 24 kk asetuksen voimaan astumisesta

Luetellaan asetuksen liitteessä III, jota komissio voi muokata tai täydentää

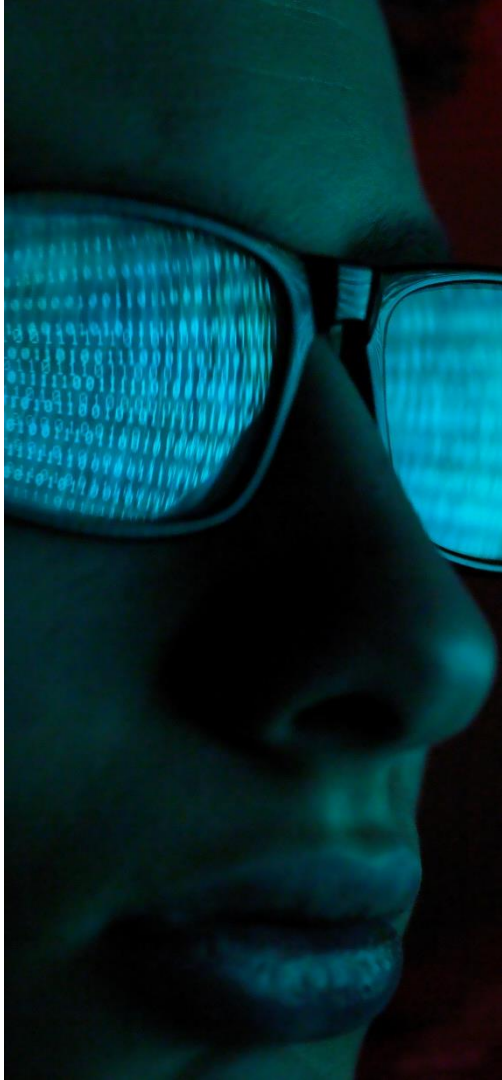
- Biometrinen etätunnistaminen ja luokittelu sekä tunteiden tunnistaminen
  - Kiellettyjen käyttötarkoitusten ulkopuolelle jäävät tapaukset
  - Ei koske yksinomaan henkilöllisyyden todentamiseen ja laitteiden avaamiseen tarkoitettuja järjestelmiä
- Kriittinen infrastruktuuri
  - Tieliikenteen ja kriittisen digitaalisen infrastruktuurin hallinnoinnin sekä veden, kaasun, lämmön ja sähkön jakelun turvakomponentit
  - Ei koske kyberturvakomponentteja
- Koulutus ja ammatillinen koulutus
  - Oppilas- ja opiskelijavalinnat
  - Oppimisen arviointi ja ohjaaminen, tarjottavan koulutuksen tason arviointi
  - Koetilanteissa kielletyn käytöksen seuranta ja havaitseminen

Reitti 2

Muut korkean riskin käyttötapaukset 2/3

- Työllisyys, työntekijöiden hallinta ja pääsy itsenäiseen ammatinharjoittamiseen
  - Rekrytoiminen ja työntekijävalinta
  - Työpaikkailmoitusten kohdentaminen
  - Työhakemusten suodattaminen ja analysointi ja hakijoiden arviointi
  - Työehtoihin ja työsuhteen päättämiseen ja ylennyksiin liittyvät päätökset
  - Työtehtävien jakaminen ja suorituksen ja käytöksen seuranta ja arviointi

- Pääsy välttämättömiin yksityisiin palveluihin ja olennaisiin julkisiin palveluihin ja etuihin
  - Ihmisen kelpoisuuden arvioiminen liittyen välttämättömiin etuisuuksiin ja palveluihin sekä näiden myöntäminen tai epääminen
  - Ihmisen luottoluokituksen arviointi, pois lukien talouspetosten havaitseminen
  - Hätäpuheluiden arviointi ja luokittelu
  - Henki- ja sairausvakuutuksen riskien arviointi ja hinnoittelu



Reitti 2

Muut korkean riskin käyttötapaukset 3/3

- Lainvalvonta
  - Rikosuhririskin, rikosriskin tai rikoksen uusimisen riskin arviointi
  - Valheenpaljastus
  - Todistusaineiston luotettavuuden arviointi
  - Rikoksen tutkintaan tai syyttämiseen liittyvä profilointi
- Maahanmuutto-, turvapaikka- ja rajavalvonnan hallinta
  - Valheenpaljastus
  - Maahantulijan turvallisuus-, terveys- ja muiden riskien arviointi
  - Turvapaikka-, viisumi- ja oleskelulupahakemusten arviointi
  - Maahantulijoiden havaitseminen tai tunnistaminen
- Oikeudenhoito ja demokraattiset prosessit
  - Tosiasioiden ja lain tutkiminen ja tulkitseminen sekä soveltaminen
  - Vaalien tai kansanäänestyksen tulokseen tai äänestäjiin vaikuttaminen

Reitti 2

Poikkeukset korkean riskin luokituksesta

Tekoälyjärjestelmää ei katsota korkeariskiseksi, jos sen tarkoitus on joku seuraavista:

1. Suorittaa kapea menettelytehtävä
   - Esim. strukturoimattoman datan strukturoiminen

Poikkeusmahdollisuus koskee vain liitteen III käyttötapauksia

2. Parantaa aiemmin suoritettua ihmisen toiminnan tulosta
   - Esim. ihmisen tuottaman tekstin tyylin muokkaaminen

3. Havaita päätöksentekomallit tai poikkeamat aiemmista päätöksentekomalleista
   - Esim. opettajan tekemän oppilasarvioinnin poikkeamien tai epäjohdonmukaisuuksien merkitseminen

4. Suorittaa valmisteleva tehtävä arviointiin, joka on olennainen korkean riskin käyttötapausten kannalta
   - Esim. tiedostojen käsitteleminen

Poikkeukset eivät päde ihmisten profilointiin tarkoitettuihin järjestelmiin

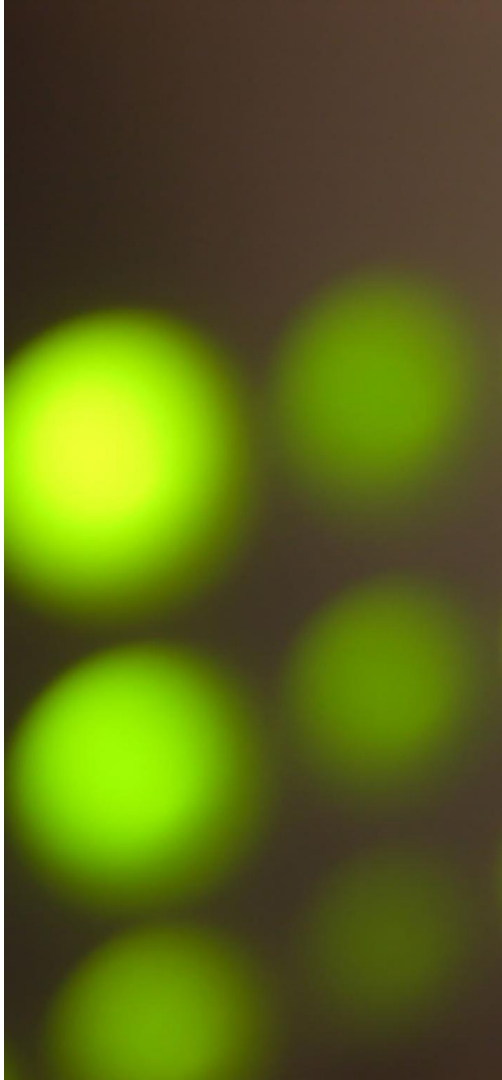
Reitti 2

Poikkeuksen dokumentointi ja rekisteröinti

- Tarjoajan, joka katsoo, että liitteessä III tarkoitettu tekoälyjärjestelmä ei ole korkeariskinen, on dokumentoitava arviointinsa ennen kuin järjestelmä saatetaan markkinoille tai otetaan käyttöön.

- Toimivaltaisten viranomaisten pyynnöstä tarjoajan on toimitettava arviointia koskevat asiakirjat.

- Tarjoajan on lisäksi noudatettava korkeariskisten järjestelmien rekisteröintivelvoitetta.
  - Ks. dia 34.

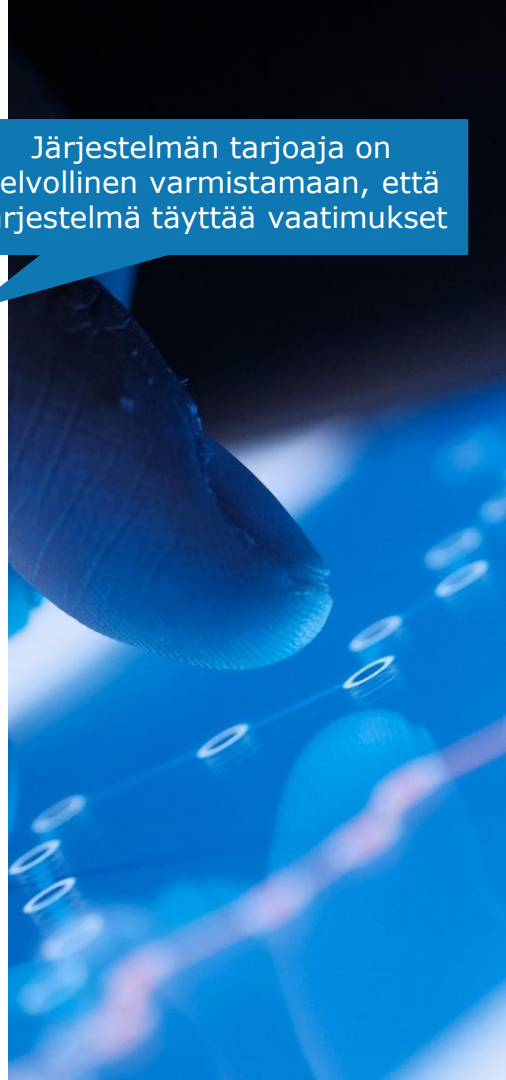


Korkean riskin tekoälyjärjestelmien vaatimukset

- Riskienhallintajärjestelmä
- Data ja datan hallinnointi
- Tekninen dokumentaatio
- Arkistointi
- Avoimuus ja käyttäjien informointi
- Ihmisen suorittama valvonta
- Tarkkuus, toimintavarmuus ja kyberturvallisuus

Järjestelmän tarjoaja on velvollinen varmistamaan, että järjestelmä täyttää vaatimukset

Vaatimusten täyttämisen tueksi on tarkoitus tuottaa harmonisoituja standardeja



Tarjoajan velvollisuus:

Vaatimustenmukaisuuden arviointi

- Ennen kuin korkean riskin tekoälyjärjestelmä saatetaan EU:n markkinoille tai muulla tavoin otetaan käyttöön, järjestelmän tarjoajan on tehtävä sille vaatimustenmukaisuuden arviointi.
  - Arviointi on toistettava, jos järjestelmää tai sen tarkoitusta muutetaan olennaisesti.

- Liitteen III piiriin tulevien korkean riskin järjestelmien vaatimuksenmukaisuuden arviointi olisi pääsääntönä voitava tehdä tarjoajan itsensä toimesta.
  - Tuoteturvasäänneltyjen tuotteiden ja turvakomponenttien (liite II) kohdalla noudatetaan lähtökohtaisesti kyseisten tuoteturvasäädösten arviointiprosesseja, mukaan lukien kolmannen osapuolen arviointia.

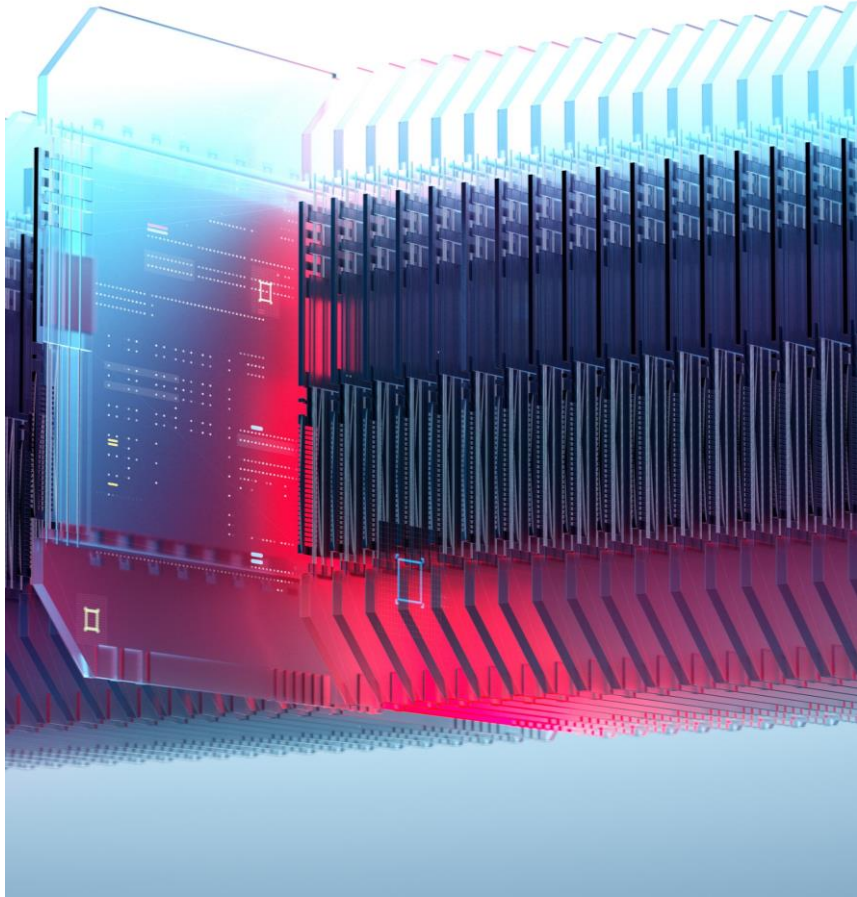
Tarjoajat voivat hyödyntää harmonisoituja standardeja osoittaakseen, että järjestelmä täyttää asetetut vaatimukset

Tarjoajan velvollisuus:

Korkean riskin järjestelmien rekisteröinti

- Liitteen III mukaisen korkean riskin tekoälyjärjestelmän tarjoajan on rekisteröitävä itsensä ja järjestelmänsä EU:n tietokantaan.
  - Tämä on tehtävä ennen järjestelmän markkinoille saattamista tai käyttöönottamista.

- Velvollisuus koskee myös liitteen III piiriin tulevaa järjestelmää, joka ei tarjoajan arvion perusteella ole korkeariskinen.



Korkean riskin järjestelmän tarjoajan muut velvollisuudet

- Laatujärjestelmä
- Markkinoille saattamisen jälkeinen seurantajärjestelmä ja -suunnitelma
- Asiakirjojen säilyttäminen (10 vuotta)
- Automaattisesti luotujen lokien säilyttäminen (6 kk)
- Vakavien vaaratilanteiden ilmoittaminen markkinavalvontaviranomaisille
- Korjaavat toimet ja tiedonantovelvollisuus
- Yhteistyö toimivaltaisten viranomaisten kanssa
- Valtuutetun EU-edustajan nimeäminen, ellei tarjoaja ole sijoittunut EU:hun

Vastuut korkean riskin tekoälyjärjestelmän arvoketjussa:

Järjestelmien muokkaaminen

- Käyttöönottaja, jakelija, maahantuoja tai muu kolmas osapuoli katsotaan korkean riskin tekoälyjärjestelmän tarjoajaksi, mikäli:
  - Laittaa nimensä tai tavaramerkkinsä korkean riskin järjestelmään.
  - Tekee merkittäviä muutoksia korkean riskin järjestelmään siten, että se pysyy korkean riskin järjestelmänä.
  - Muuttaa tekoälyjärjestelmän käyttötarkoitusta, jota ei ole luokiteltu korkeariskiseksi, siten, että järjestelmästä tulee korkean riskin järjestelmä.

Näissä tilanteissa tarjoajaa, joka alun perin saattoi järjestelmän markkinoille tai otti sen käyttöön, ei enää pidetä kyseisen järjestelmän tarjoajana

Alkuperäisen tarjoajan on toimittava yhteistyössä ja asetettava saataville tarvittavat tiedot sekä annettava kohtuullinen tekninen pääsy ja muu apu, jota vaaditaan velvoitteiden täyttämiseksi, ellei tämä ole nimenomaisesti sulkenut pois järjestelmänsä muuttamisen korkean riskin järjestelmäksi

Vastuut korkean riskin tekoälyjärjestelmän arvoketjussa:

Järjestelmien työkalujen, palveluiden ja komponenttien toimittajat

- Kolmannet osapuolet, jotka toimittavat korkean riskin järjestelmän tarjoajalle järjestämään liittyviä työkaluja, palveluita ja komponentteja, on sovittava kirjallisesti toimitettavista tiedoista ja tuesta, jotta tarjoaja voi noudattaa asetuksen velvollisuuksia.
  - Ei koske ilmaisin, avoimin lisenssein tarjottuja työkaluja, palveluita ja komponentteja.

Komission tekoälytoimisto voi kehittää ja suositella vapaaehtoisia mallisopimusehtoja korkean riskin tekoälyjärjestelmien tarjoajien ja kolmansien osapuolten välillä

Korkean riskin järjestelmän käyttöönottajan velvollisuudet

- Käytettävä ja valvottava järjestelmää käyttöohjeiden mukaisesti.
- Osoitettava järjestelmän valvonta henkilöille, joilla on tarvittava pätevyys, koulutus ja valtuudet sekä tarvittava tuki.
- Varmistettava, että syöttötiedot ovat merkityksellisiä ja riittävän edustavia järjestelmän käyttötarkoituksen kannalta – siinä määrin kuin käyttöönottaja hallitsee syöttödataa.
- Informoitava järjestelmän tarjoajaa ja valvontaviranomaista havaitsemistaan riskeistä.
- Säilytettävä järjestelmän automaattisesti luomat lokit vähintään 6 kk ajan – siltä osin kuin ne ovat käyttöönottajan hallinnassa.
- Informoitava työntekijöitä työpaikalla käyttöön otettavasta järjestelmästä.
- Informoitava luonnollisia henkilöitä, jotka ovat korkeariskisen päätöksiä tekevän tai niissä avustavan järjestelmän kohteina.
- Julkisen sektorin käyttöönottajien on lisäksi rekisteröidyttävä EU:n tietokantaan

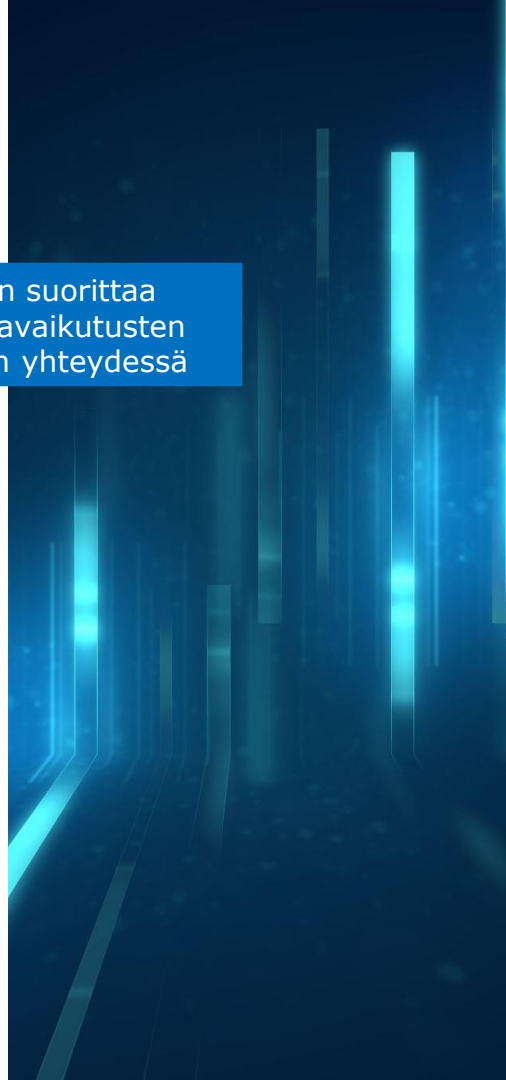
Julkisen sektorin käyttöönottajan velvollisuus:

Perusoikeusvaikutusten arviointi

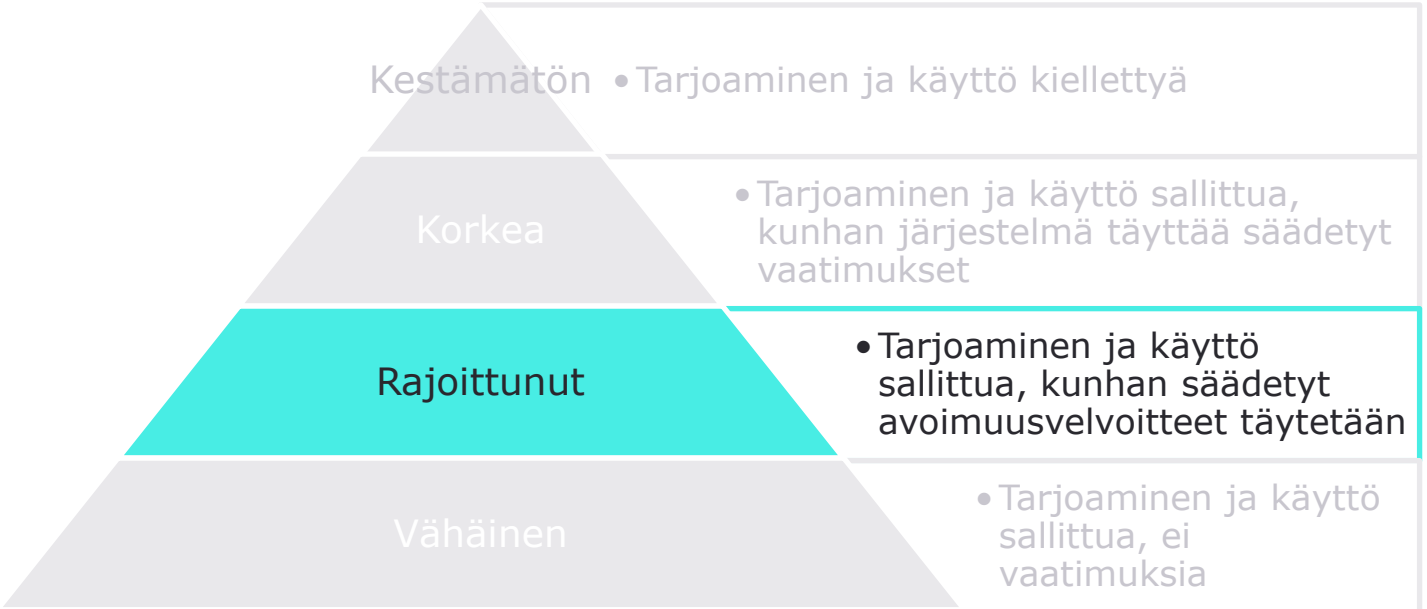
Voidaan suorittaa tietosuojavaikutusten arvioinnin yhteydessä

- Korkean riskin tekoälyjärjestelmien käyttöönottajien on suoritettava perusoikeusvaikutusten arviointi ennen järjestelmän käyttöönottoa.
  - Velvoite koske julkisen sektorin elimiä sekä keskeisiä julkisia palveluita tuottavia toimijoita.

- Arvioinnin tulokset on ilmoitettava toimivaltaiselle viranomaiselle.



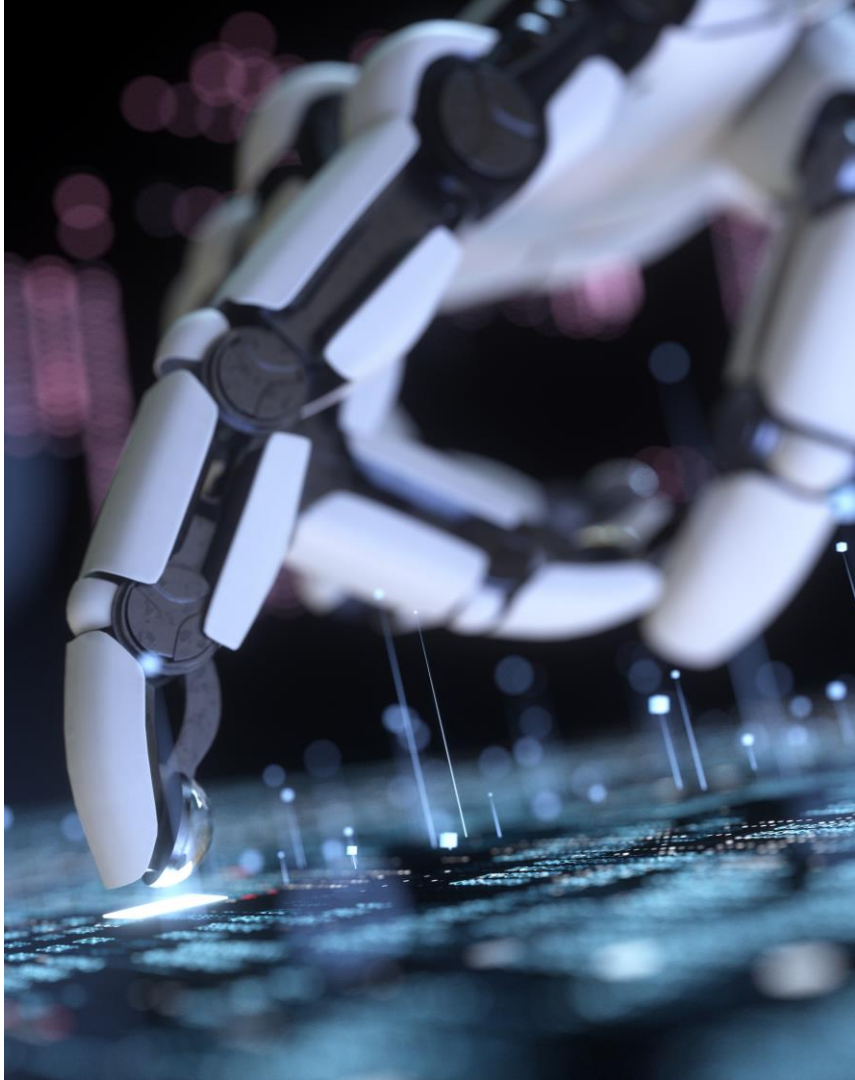
Rajoittuneen riskin käyttötapaukset ja avoimuusvelvoitteet



Järjestelmäriskejä sisältävät mallit	• Lisävaatimukset
Kaikki yleiskäyttöiset mallit	• Avoimuus- ja informointivaatimukset

Ihmisen kanssa vuorovaikuttavat tekoälyjärjestelmät

- Tarjoajien on varmistettava, että tekoälyjärjestelmät, jotka on tarkoitettu vuorovaikuttamaan välittömästi ihmisten kanssa, toteutetaan siten, että kyseiset henkilöt saavat tiedon heidän asioivan tekoälyjärjestelmän kanssa.
  - Ei koske järjestelmiä, jotka on lailla valtuutettu rikoksentorjuntatarkoituksiin.



Generatiiviset tekoälyjärjestelmät:

Synteettisten tuotosten merkitseminen

- Synteettistä ääni-, kuva-, video- tai tekstisisältöä tuottavien tekoälyjärjestelmien tarjoajien on varmistettava, että mallin tai järjestelmän tuotokset on merkitty koneellisesti luettavassa muodossa ja havaittavissa keinotekoisiksi tai manipuloiduiksi.
  - Käytettyjen tekniset ratkaisujen on oltava tehokkaita, yhteentoimivia, kestäviä ja luotettavia siinä määrin kuin se on teknisesti mahdollista, ottaen huomioon erityyppisten sisältöjen erityispiirteet, toteutuskustannukset ja yleisesti tunnustettu tekniikan taso.
  - Ei koske järjestelmiä, jotka suorittavat sisällönmuokkauksen aputoimintoa tai jotka eivät olennaisesti muuta käyttöönottajan toimittamaa syöttödataa tai sen semantiikkaa.

Generatiiviset tekoälyjärjestelmät:

Deep fake -väärennösten avoimuus

- Tekoälyjärjestelmän käyttöönottajien, jotka luovat tai manipuloivat kuva-, ääni- tai videosisältöä, joka muodostaa syväväärennöksen (”deep fake”), on ilmoitettava, että sisältö on keinotekoisesti tuotettua.
  - Jos syväväärennös on osa taiteellista, luovaa, satiirista tai fiktiivistä teosta, avoimuusvelvoite rajoittuu tällaisen sisällön paljastamiseen tavalla, joka ei estä sen näyttämistä tai siitä nauttimista.
  - Ei koske lailla valtuutettua, rikoksentorjuntatarkoituksiin tarkoitettua käyttöä.

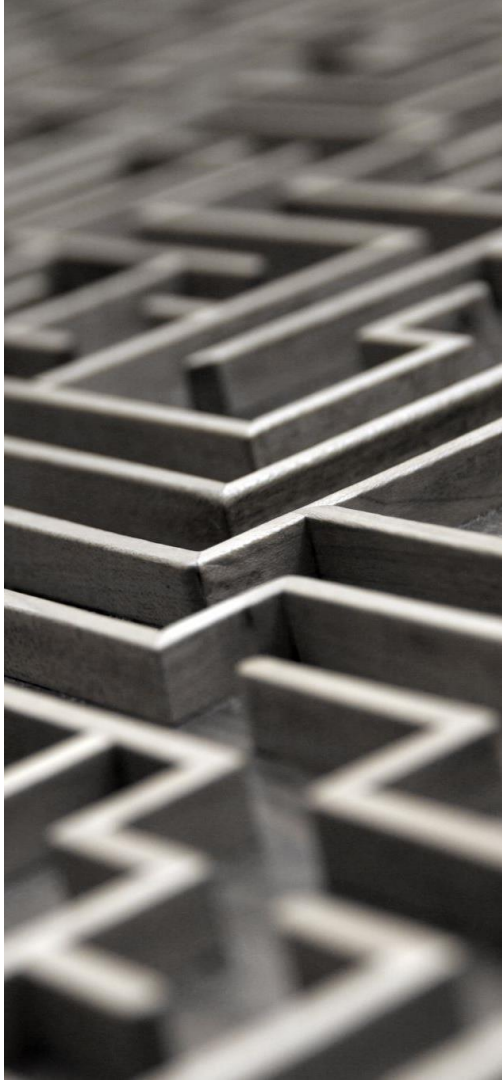
Generatiiviset tekoälyjärjestelmät:

Yleistä etua koskevan tekstin avoimuus

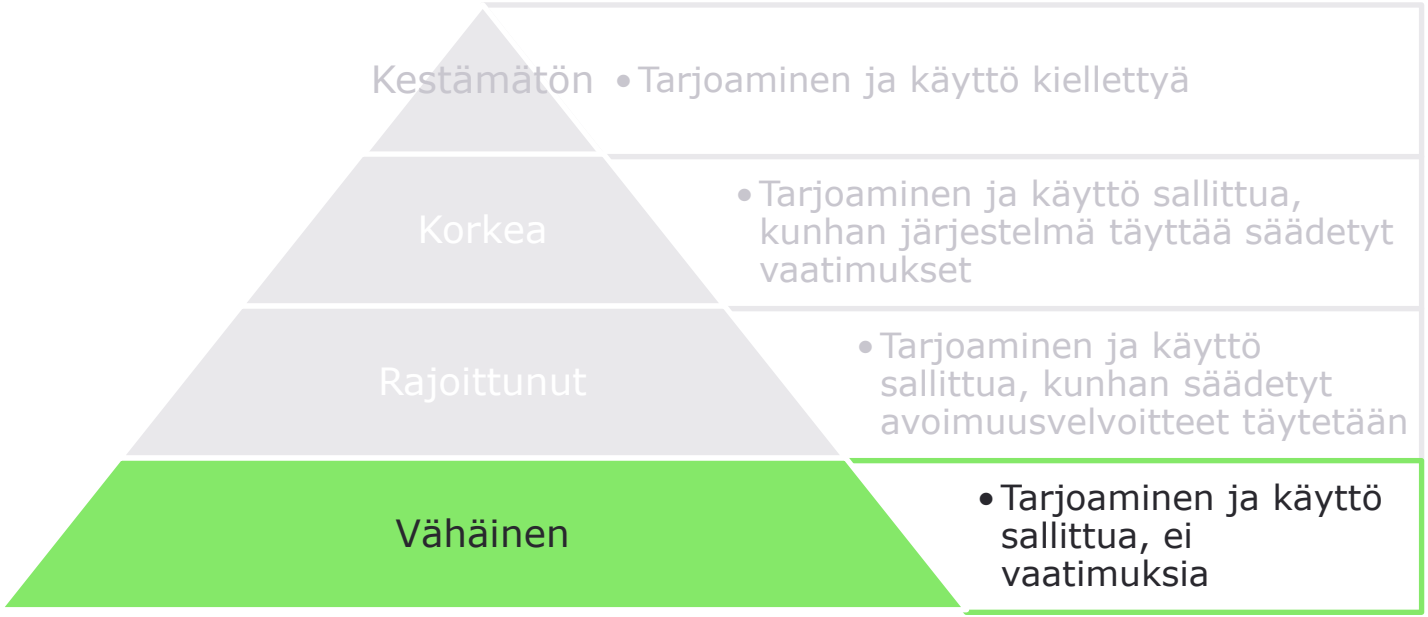
- Tekoälyjärjestelmän käyttöönottajien, jotka luovat tai manipuloivat tekstiä, joka on julkaistu tarkoituksena tiedottaa yleisölle yleistä etua koskevista asioista, on ilmoitettava, että teksti on tuotettu keinotekoisesti.
  - Ei koske tilanteita, joissa järjestelmän luoma sisältö on käynyt läpi ihmisen tarkastelun tai toimituksellisen valvonnan ja luonnollisella henkilöllä tai oikeushenkilöllä on toimituksellinen vastuu sisällön julkaisemisesta.
  - Ei koske lailla valtuutettua, rikoksentorjuntatarkoituksiin tarkoitettua käyttöä.

Tunteidentunnistusjärjestelmät ja biometriset luokittelujärjestelmät

- Tunteidentunnistusjärjestelmän tai biometrisen luokitusjärjestelmän käyttöönottajien on tiedotettava järjestelmän toiminnasta niille altistuneita luonnollisia henkilöitä.
  - Ei koske lailla valtuutettua, rikoksentorjuntatarkoituksiin tarkoitettua käyttöä.



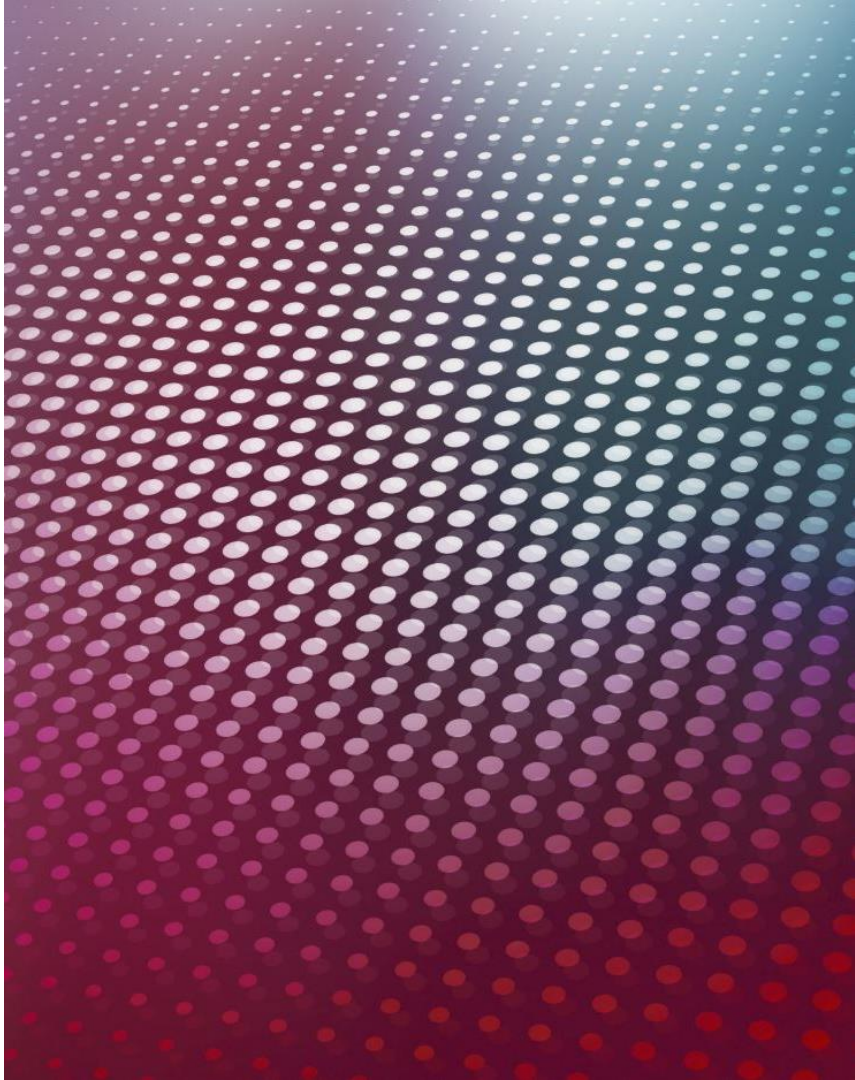
Vähäisen riskin käyttötapaukset ja vapaaehtoiset käytännesäännöt



Järjestelmäriskejä sisältävät mallit	• Lisävaatimukset
Kaikki yleiskäyttöiset mallit	• Avoimuus- ja informointivaatimukset

Vapaehtoiset käytännesäännöt

- Vähäisen riskin tekoälyjärjestelmien tarjoajat voivat varmistaa järjestelmänsä luotettavuuden kehittämällä omia vapaaehtoisia käytännesääntöjä tai noudattamalla muiden edustavien tahojen hyväksymiä käytännesääntöjä.



Yleiskäyttöisten tekoälymallien vaatimukset

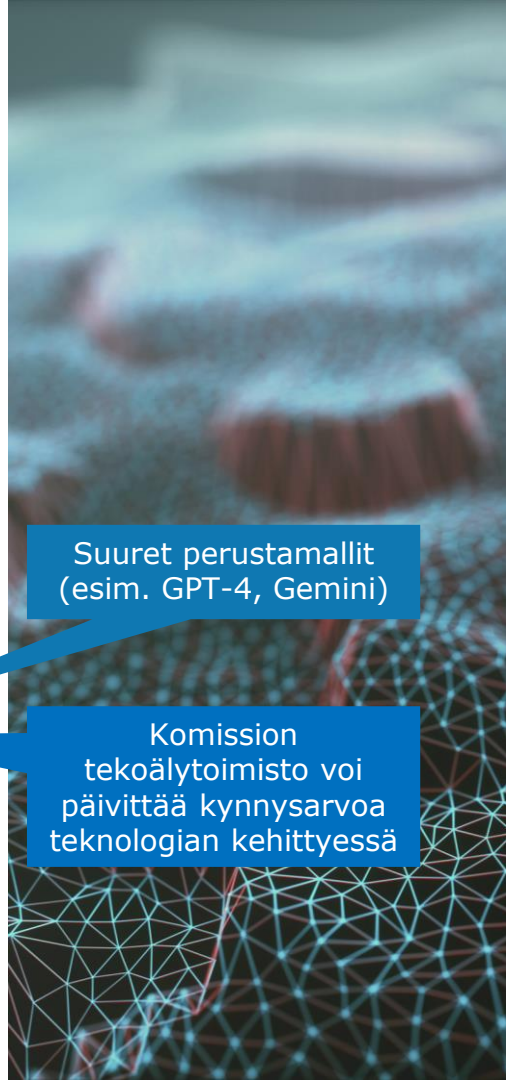
Järjestelmäriskejä sisältävät mallit	• Lisävaatimukset
Kaikki yleiskäyttöiset mallit	• Avoimuus- ja informointivaatimukset

Yleiskäyttöiset tekoälymallit:

Kaksitasoinen riskiluokittelu

1. Kaikki yleiskäyttöiset tekoälymallit
   - Ks. määritelmä dialta 13.

2. Yleiskäyttöiset tekoälymallit, joiden katsotaan sisältävän järjestelmäriskejä
   - Malli kuuluu tähän luokkaan,
     a) jos sen koulutuksessa on käytetty laskentatehoa vähintään 10^25 FLOPs
     b) TAI jos Euroopan tekoälytoimisto katsoo mallin muihin tekijöihin perustuen käsittävän systeemisiä riskejä (esim. parametrimäärä, käyttäjämäärä, autonomisuuden aste).



Kaikkia yleiskäyttöisiä malleja koskevat vaatimukset

- Dokumentaatio ja arviointi:
  - Laatia ja pitää ajan tasalla mallin tekniset dokumentaatiot mukaan lukien sen koulutus- ja testausprosessi ja sen arvioinnin tulokset.
- Jatkotarjoajan informointi:
  - Laatia, pitää ajan tasalla ja asettaa saataville tiedot ja dokumentaatio tekoälyjärjestelmien tarjoajille, jotka aikovat integroida yleiskäyttöisen tekoälymallin tekoälyjärjestelmäänsä.
  - Tämän on annettava tekoälyjärjestelmien tarjoajille hyvä käsitys mallin ominaisuuksista ja rajoituksista, jotta nämä voivat noudattaa asetuksen velvoitteita.
- Tekijänoikeudet:
  - Ottaa käyttöön käytäntö EU:n tekijänoikeusääntelyn kunnioittamiseksi.
  - Laatia ja asettaa julkisesti saataville riittävän yksityiskohtainen tiivistelmä sisällöistä, joita on käytetty tekoälymallin koulutukseen.

Sovelletaan 12 kk asetuksen voimaan astumisesta

Immateriaalioikeuksia ja liikesalaisuuksia kunnioittaen ja suojaten

Avoimen lähdekoodin malleja, jotka eivät ole järjestelmäriskisiä, koskevat vain tekijänoikeusvaatimukset

Järjestelmäriskisten mallien lisävaatimukset

- Mallin arviointi standardoitujen protokollien ja työkalujen mukaisesti, sisältäen mallin vastatestauksen (”red teaming”) suorittamisen ja dokumentoinnin
- Järjestelmäriskien arviointi ja lieventäminen
- Vaaratilanteiden dokumentointi ja raportointi
- Kyberturvallisuuden varmistaminen mallille ja sen fyysiselle infrastruktuurille

Sovelletaan 12 kk asetuksen voimaan astumisesta

Vaatimukset täyttääkseen mallien tarjoajat voivat nojata alkuvaiheessa yleisiin käytännesääntöihin ja myöhemmin harmonisoituun standardiin

Lisävaatimukset koskevat myös avoimen lähdekoodin malleja, jotka ovat järjestelmäriskisiä

Avoimen lähdekoodin yleiskäyttöiset mallit

- Ilmaisella ja avoimen lähdekoodin lisenssillä tarjotun yleiskäyttöisen tekoälymallin, jonka parametrit ja käyttöinformaatio ovat julkisesti saatavilla, on noudatettava yleiskäyttöisten mallien tekijänoikeuksia koskevia vaatimuksia.

- Jos avoimen lähdekoodin mallin katsotaan sisältävän järjestelmäriskejä, on sen noudatettava kaikkia vaatimuksia.

Innovointia edistävät toimet

Tekoälyn sääntelyhiekkalaatikko

- Jäsenmaan toimivaltaisen viranomaisen on perustettava yksin tai muiden jäsenmaiden kanssa vähintään yksi riittävästi resursoitu kansallisen tason sääntelyhiekkalaatikko.

- Hiekkalaatikko tarjoaa tekoälyjärjestelmien tarjoajille mahdollisuuden kehittää, kouluttaa, validoida ja testata innovatiivisia tekoälyjärjestelmiä ennen niiden markkinoille saattamista tai käyttöönottoa.
  - Hiekkalaatikointi on mahdollista rajatun ajan ja toimivaltaisten viranomaisten kanssa sovitun suunnitelman mukaisesti.
  - Viranomaisten on tarjottava ohjeistusta ja tukea asetuksen vaatimuksista ja velvoitteista ja niiden täyttämisestä.

Oltava toiminnassa, kun säädöksen voimaan astumisesta on kulunut 24 kk

Maksutonta pk-yrityksille ja startupeille – poikkeuksellisia kustannuksia voidaan periä

Hiekkalaatikointiin osallistuvat tarjoajat eivät joudu hallinnollisten sakkojen kohteeksi asetuksen rikkomisesta, mutta ovat vastuussa kolmansille osapuolille aiheutuneista vahingoista

Hiekkalaatikointi ja asetuksen vaatimusten täyttäminen

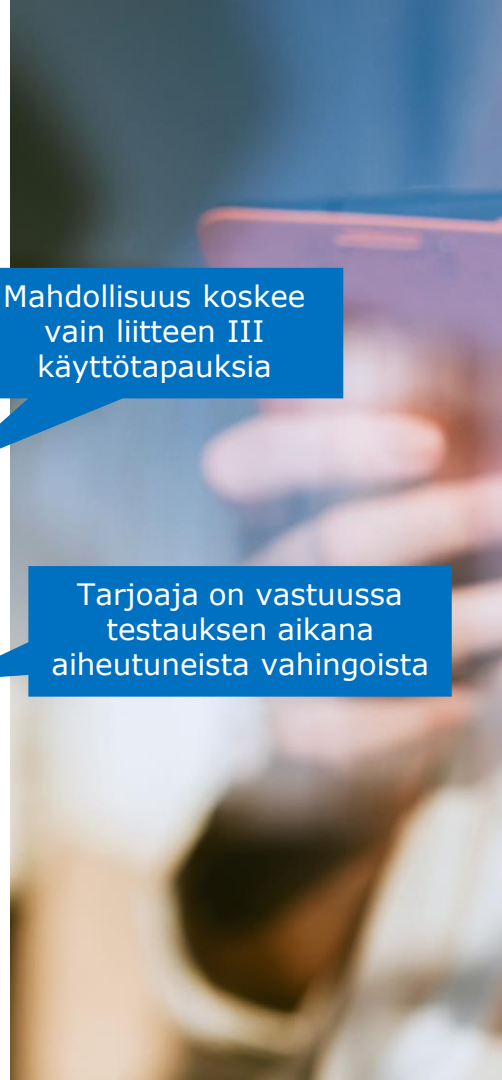
- Toimivaltaisen viranomaisen on toimitettava tekoälyjärjestelmän tarjoajan pyynnöstä kirjallinen todiste hiekkalaatikossa onnistuneesti suoritetuista toimista.

- Toimivaltaisen viranomaisen on myös toimitettava raportti, jossa esitetään yksityiskohtaisesti hiekkalaatikossa suoritetut toimet sekä niihin liittyvät tulokset ja oppimistulokset.

- Tarjoajat voivat käyttää näitä asiakirjoja osoittaakseen, että he noudattavat asetuksen vaatimuksia ja velvoitteita.
  - Valvontaviranomaisten ja ilmoitettujen arviointilaitosten on otettava myönteisesti huomioon nämä asiakirjat nopeuttaakseen vaatimustenmukaisuuden arviointimenettelyjä.

Korkean riskin tekoälyjärjestelmän testaaminen todellisissa olosuhteissa

- Korkean riskin tekoälyjärjestelmän tarjoaja voi suorittaa järjestelmän testauksen todellisissa olosuhteissa tekoälyn sääntelyhiekkalaatikoiden ulkopuolella enintään 12 kk ajan.
  - Tarjoajan on noudatettava tästä asetuksessa säädettyjä määräyksiä, mukaan lukien niiden mukaisesti laadittavaa, markkinavalvontaviranomaisen hyväksymää testaussuunnitelmaa.
  - Testaukseen osallistuvilta henkilöiltä on saatava tietoinen suostumus.



Muita innovaatioita edistäviä toimia

- Jäsenmaat:
  - Tehtävä tietoisuutta kasvattavia ja koulutuksellisia toimia säädöksen soveltamisesta erityisesti pk-yrityksille, startupeille ja viranomaisille.
  - Tarjottava pk-yrityksille, startupeille ja viranomaisille neuvoja ja vastattava näiden tiedusteluihin liittyen säädöksen toimeenpanoon.
  - Edistettävä pk-yritysten ja muiden sidosryhmien osallistumista standardointityöhön.
- Komissio:
  - Tarjottava standardoituja malleja asetuksen soveltamisalaan kuuluville aloille.
  - Ylläpidettävä alustaa, joka tarjoaa helppokäyttöistä asetukseen liittyvää tietoa.
  - Järjestettävä tiedotuskampanjoita.

Hallinto, valvonta ja rangaistukset

Hallinto ja valvonta:

Jäsenmaataso

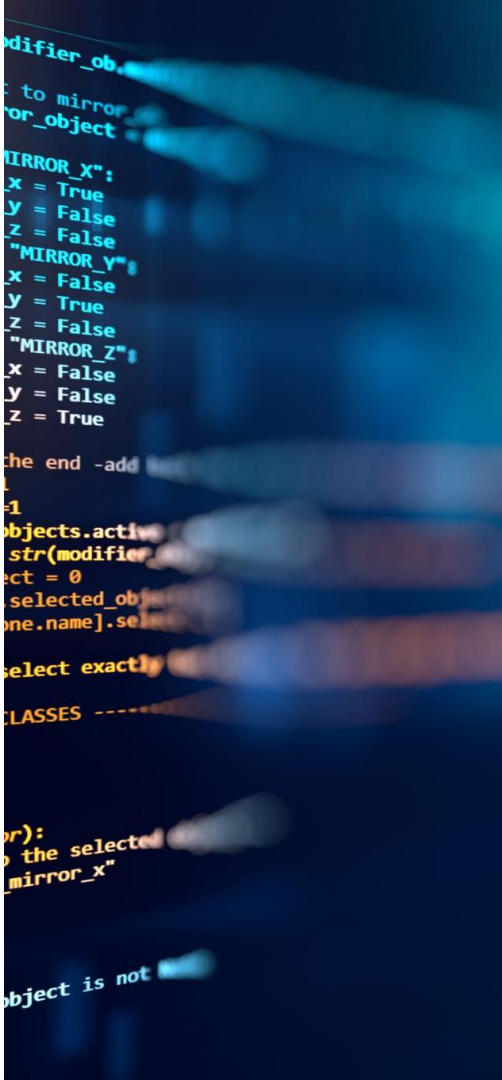
- Jäsenvaltion on perustettava tai nimettävä vähintään yksi ilmoittamisesta vastaava viranomainen ja vähintään yksi markkinavalvontaviranomainen kansallisiksi toimivaltaisiksi viranomaisiksi.

Hallinto ja valvonta:

EU-taso

- Euroopan tekoälylautakunta:
  - Edustettuina jäsenmaiden valvontaviranomaiset, komissio ja Euroopan tietosuojavaltuutettu.
  - Neuvoo ja avustaa komissiota ja jäsenvaltioita ja edistää näin asetuksen yhdenmukaista ja tehokasta soveltamista.
  - Teknistä lisäasiantuntemusta lautakunnalle tarjoaa neuvoa-antava foorumi, joka edustaa teollisuutta, startup- ja pk-yrityksiä, kansalaisyhteiskuntaa ja tiedeyhteisöä.

- Euroopan tekoälytoimisto:
  - Perustetaan komission yhteyteen.
  - Luokittelee ja valvoo yleiskäyttöisiä tekoälymalleja ja tekee yhteistyötä tekoälylautakunnan kanssa.
  - Toimistoa tukee riippumattomista asiantuntijoista koostuva tieteellinen paneeli.



Komission ohjeistukset

Komissio on velvollinen tuottamaan seuraavat ohjeistukset säädöksen käytännön täytäntöönpanosta:

- Korkean riskin järjestelmien vaatimukset ja arvoketjun velvollisuudet
- Kielletyt käyttötapaukset
- Järjestelmien olennaisiin muokkauksiin liittyvät säännökset
- Rajoittuneen riskin järjestelmien ja mallien avoimuusvelvoitteet
- Tekoälyasetuksen suhde tuoteturvasääntelyyn (liite II) ja muuhun sääntelyyn
- Tekoälyjärjestelmän määritelmän soveltaminen

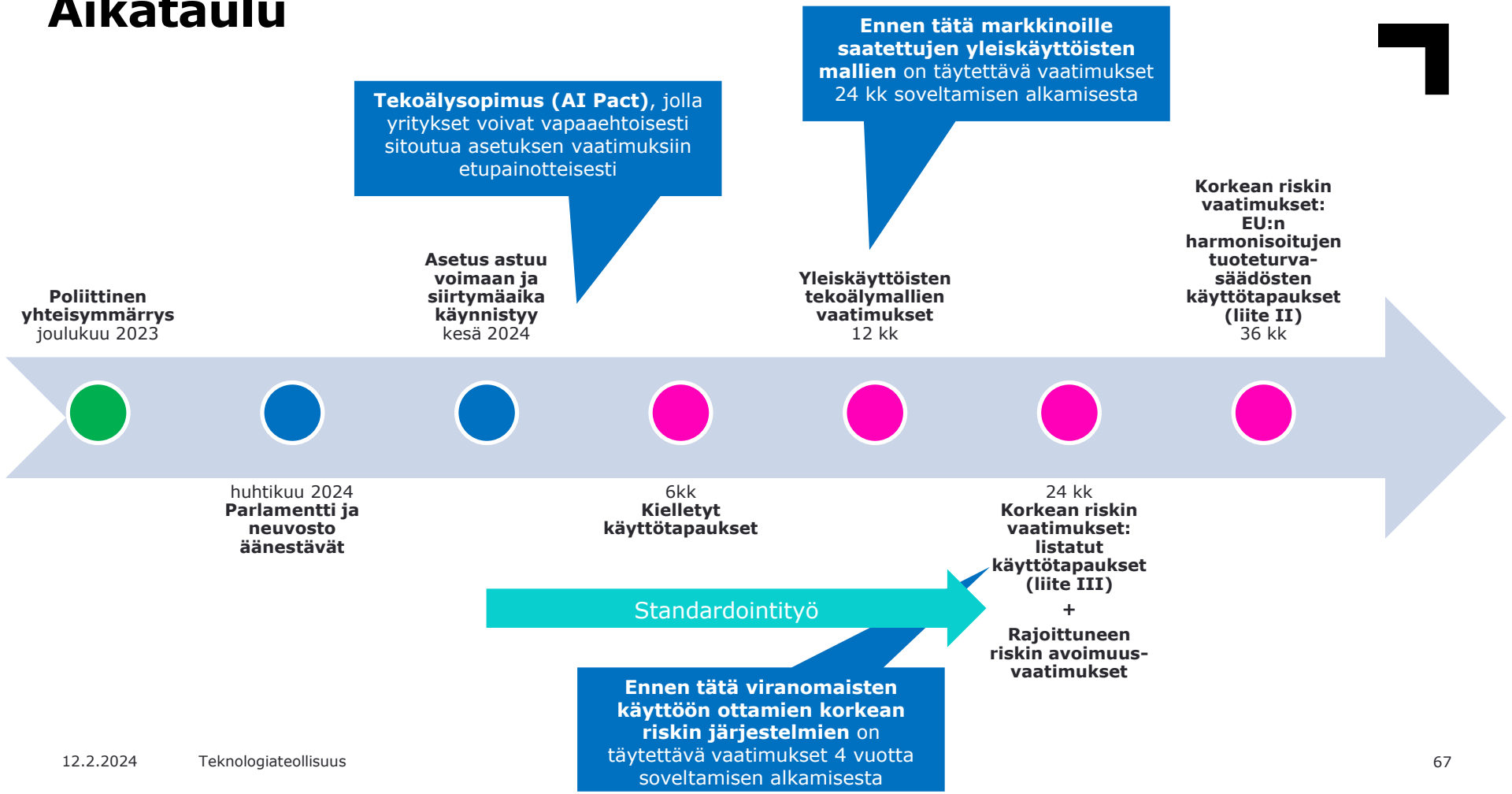
Rangaistukset

- Kiellettyjen käyttötapausten rikkominen:
  - Enintään 35 miljoona euroa tai 7 prosenttia edellisen tilikauden maailmanlaajuisesta vuotuisesta kokonaisliikevaihdosta.
- Yleiskäyttöisten tekoälymallien sääntöjen ja asetuksen muiden vaatimusten tai velvoitteiden rikkominen:
  - Enintään 15 miljoonaa euroa tai 3 prosenttia liikevaihdosta.
- Epätäydellisen tai harhaanjohtavan tiedon toimittaminen:
  - Enintään 7,5 miljoonaa euroa tai 1,5 prosenttia liikevaihdosta.

Jokaisessa rikkomusluokassa enimmäiskynnys on kahdesta summasta pienempi pk-yritysten osalta ja korkeampi muiden yritysten osalta

Soveltamisaikataulu

Aikataulu


Tekoälysopimus (AI Pact), jolla yritykset voivat vapaaehtoisesti sitoutua asetuksen vaatimuksiin etupainotteisesti
Ennen tätä markkinoille saatettujen yleiskäyttöisten mallien on täytettävä vaatimukset 24 kk soveltamisen alkamisesta
Poliittinen yhteisymmärrys
joulukuu 2023
huhtikuu 2024
Parlamentti ja neuvosto äänestävät
Asetus astuu voimaan ja siirtymäaika käynnistyy
kesä 2024
6kk
Kielletyt käyttötapaukset
Yleiskäyttöisten tekoälymallien vaatimukset
12 kk
24 kk
Korkean riskin vaatimukset: listatut käyttötapaukset (liite III)
+
Rajoittuneen riskin avoimuus-vaatimukset
Korkean riskin vaatimukset: EU:n harmonisoitujen tuoteturva-säädösten käyttötapaukset (liite II)
36 kk
Standardointityö
Ennen tätä viranomaisten käyttöön ottamien korkean riskin järjestelmien on täytettävä vaatimukset 4 vuotta soveltamisen alkamisesta