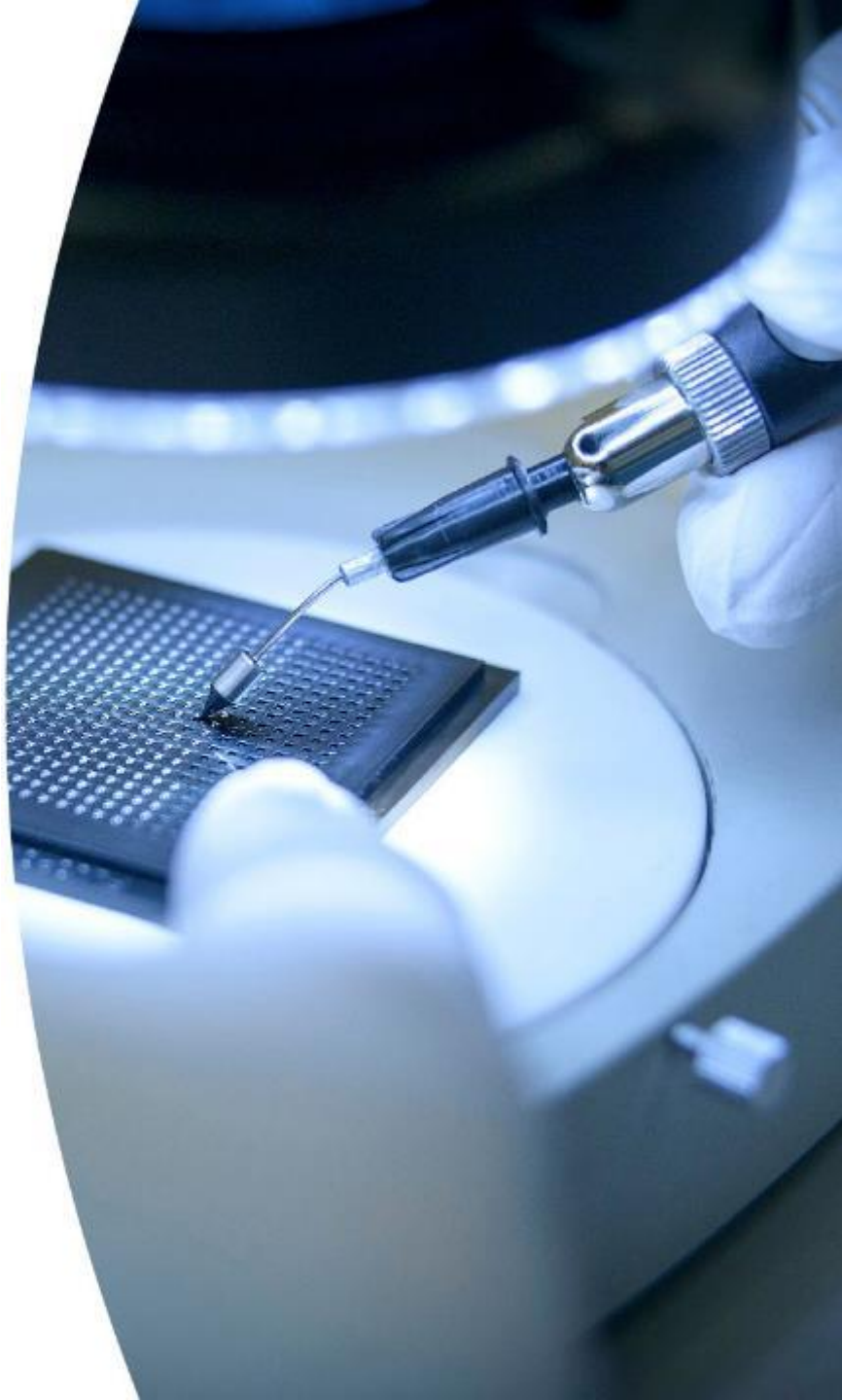
VALTIOVARAINMINISTERIÖ
FINANSMINISTERIET

EU:n tekoälyasetus ja kansallinen APT-yleissääntely

Antti Helin
25.3.2022

Tekoälyasetuksen rakenne, soveltamisala ja keskeinen sääntely



Tekoälyasetuksen rakenne

I osasto: Yleiset säännökset	Artikla 1–4
II osasto: Kielletyt tekoälyyn liittyvät käytännöt	Artikla 5
III osasto: Suuririskiset tekoälyjärjestelmät	Artikla 6–51
IV osasto: Tiettyjä tekoälyjärjestelmiä koskevat läpinäkyvyysvelvoitteet	Artikla 52
V osasto: Innovointia tukevat toimenpiteet	Artikla 53–55
VI osasto: Hallinnointi	Artikla 56–59
VII osasto: Erillisiä suuririskisiä tekoälyjärjestelmiä koskeva EU:n tietokanta	Artikla 60
VIII osasto: Markkinoille saattamisen jälkeinen seuranta, tietojen jakaminen, markkinavalvonta	Artikla 61–68
IX osasto: Käytännesäännöt	Artikla 69
X osasto: Luottamuksellisuus ja seuraamukset	Artikla 70–72
XI osasto: Säädösvallan siirto ja komiteamenettely	Artikla 73-85
Liitteet	Liite I–IX

Tekoälyasetuksen rakenne

I osasto: Yleiset säännökset	Artikla 1–4
II osasto: Kielletyt tekoälyyn liittyvät käytännöt	Artikla 5
III osasto: Suuririskiset tekoälyjärjestelmät	Artikla 6–51
IV osasto: Tiettyjä tekoälyjärjestelmiä koskevat läpinäkyvyysvelvoitteet	Artikla 52
~~V osasto: Innovointia tukevat toimenpiteet~~	~~Artikla 53–55~~
~~VI osasto: Hallinnointi~~	~~Artikla 56–59~~
~~VII osasto: Erillisiä suuririskisiä tekoälyjärjestelmiä koskeva EU:n tietokanta~~	~~Artikla 60~~
~~VIII osasto: Markkinoille saattamisen jälkeinen seuranta, tietojen jakaminen, markkinavalvonta~~	~~Artikla 61–68~~
~~IX osasto: Käytännesäännöt~~	~~Artikla 69~~
~~X osasto: Luottamuksellisuus ja seuraamukset~~	~~Artikla 70–72~~
~~XI osasto: Säädösvallan siirto ja komiteamenettely~~	~~Artikla 73-85~~
Liitteet	Liite I–IX

Tekoälyjärjestelmän määritelmä (artikla 3(1) ja liite I)

‘artificial intelligence system’ (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;

Liite I, jota voidaan muuttaa delegoidulla asetuksella:

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

(c) Statistical approaches, Bayesian estimation, search and optimization methods

Asetuksen vaatimukset kohdistuvat tuotanto- ja käyttöketjuun vaatimuksesta riippuen (artikla 2)

Tätä asetusta sovelletaan

(a) tekoälyjärjestelmien *tarjoajiin*, jotka saattavat markkinoille tai ottavat käyttöön tekoälyjärjestelmiä unionissa, riippumatta siitä, ovatko kyseiset tarjoajat sijoittautuneet unioniin vai kolmanteen maahan;

(b) unionissa sijaitseviin tekoälyjärjestelmien *käyttäjiin*;

(c) kolmannessa maassa sijaitseviin tekoälyjärjestelmien tarjoajiin ja käyttäjiin, kun järjestelmän tuottamaa tulosta käytetään unionissa.

Tuotanto- ja käyttöketjun toimijat (artikla 3)

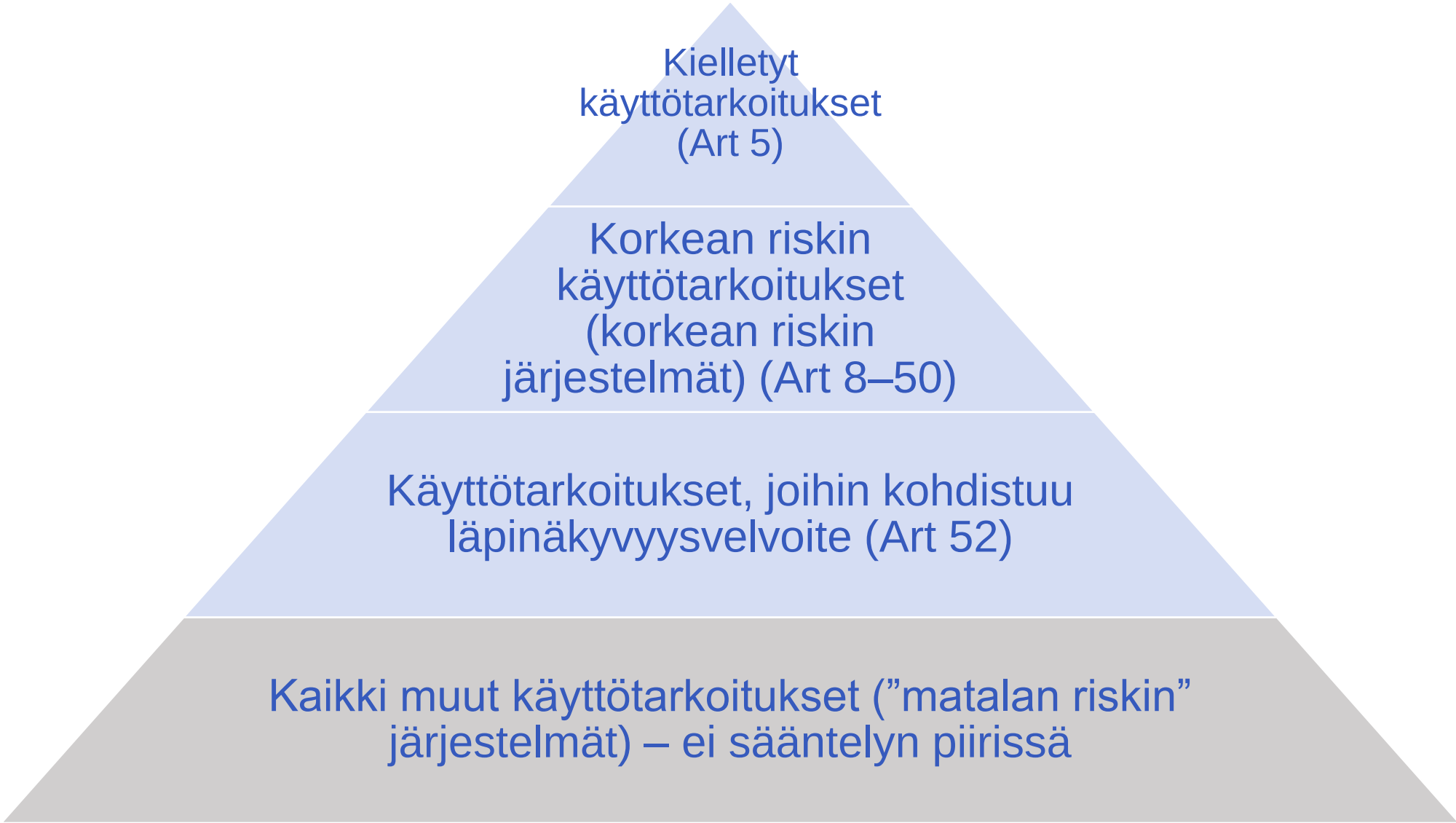
’tarjoajalla’ luonnollista tai oikeushenkilöä, viranomaista, virastoa tai muuta tahoa, joka kehittää tai kehityttää tekoälyjärjestelmän sen saattamiseksi markkinoille tai ottamiseksi käyttöön omalla nimellään tai tavaramerkillään joko maksua vastaan tai maksutta;

’käyttäjällä’ luonnollista tai oikeushenkilöä, viranomaista, virastoa tai muuta tahoa, joka käyttää valvonnassaan olevaa tekoälyjärjestelmää, paitsi jos tekoälyjärjestelmää käytetään henkilökohtaisessa muussa kuin ammattitoiminnassa;

’maahantuojalla’ unioniin sijoittautunutta luonnollista tai oikeushenkilöä, joka saattaa markkinoille tai ottaa käyttöön tekoälyjärjestelmän, jolla on unionin ulkopuolelle sijoittautuneen luonnollisen tai oikeushenkilön nimi tai tavaramerkki;

’jakelijalla’ muuta toimitusketjuun kuuluvaa luonnollista tai oikeushenkilöä kuin tarjoajaa tai maahantuojaa, joka asettaa tekoälyjärjestelmän saataville unionin markkinoilla vaikuttamatta sen ominaisuuksiin;

’toimijalla’ tarjoajaa, käyttäjää, valtuutettua edustajaa, maahantuojaa ja jakelijaa;


Kielletyt käyttötarkoitukset (Art 5)
Korkean riskin käyttötarkoitukset (korkean riskin järjestelmät) (Art 8–50)
Käyttötarkoitukset, joihin kohdistuu läpinäkyvyysvelvoite (Art 52)
Kaikki muut käyttötarkoitukset (”matalan riskin” järjestelmät) – ei sääntelyn piirissä

Kielletyt tekoälyn käyttötarkoitukset (artikla 5)

- Alitajuisesti (subliminal) henkilön käytökseen vaikuttaminen haitallisesti tai tavalla joka on omiaan aiheuttamaan haittaa.
- Haavoittuvassa asemassa olevien ihmisryhmien hyväksikäyttäminen tavalla joka aiheuttaa tai on omiaan aiheuttamaan haittaa.
- Henkilöiden luotettavuuden tai käytöksen pisteyttäminen (social scoring) julkisen hallinnon toimesta, jos pisteyttämistä käytetään ihmisten tai ihmisryhmien eriarvoiseen kohtelemiseen ja tietoja käytetään irrallaan alkuperäisestä asiayhteydestään.
- Reaaliaikainen etänä tapahtuva biometrinen tunnistaminen ilman painavaa, lueteltua syytä; jos käytetään, käyttöön kohdistuu erilaisia rajoituksia.

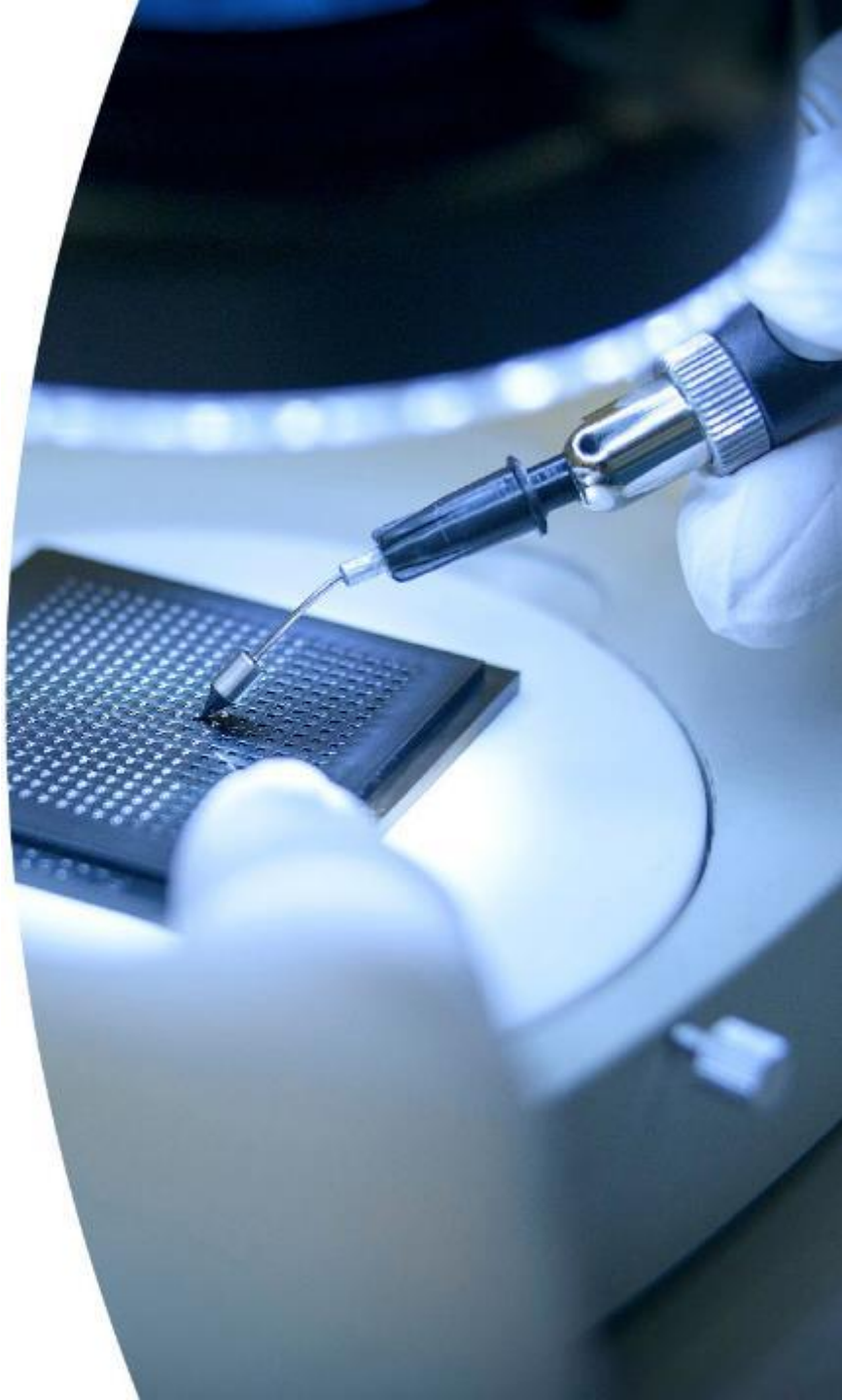
Korkean riskin tekoälyjärjestelmät (Artiklat 6-51)

- Tekoälyjärjestelmät, joita käytetään laitteiden turvallisuuskomponenttina sektoreilla, joita säädellään liitteessä II mainituilla EU-säädöksillä:
  - Ilmailu, raide- ja tieliikenne
  - Maa- ja metsätalouskoneet
  - Merenkulku
- Tekoälyjärjestelmät, joita käytetään liitteen III mainitsemissa tarkoituksissa:
  - Etänä tapahtuva biometrinen tunnistaminen (ei vain reaaliaikainen)
  - Kriittisen infrastruktuurin ylläpito ja hallinta
  - Koulutukseen sisään ottaminen ja opintosuoritusten arviointi
  - Työhön ottaminen ja irtisanominen
  - Etuuspäätökset
  - Luottokelpoisuus
  - Ensiavun lähettäminen ja priorisointi
  - Lainvalvonta ja –käyttö
  - Maahanmuutto

Tiettyjä tekoälyjärjestelmiä koskeva läpinäkyvyys (artikla 52)

- Ihmisten kanssa toimivien järjestelmien on kerrottava ihmiselle, että kyseessä on tekoälyjärjestelmä, tai ne on suunniteltava niin, että ihminen ymmärtää tämän, ellei se muuten ole ilmeistä.
- Tunteita tunnistavien ja biometrisiä tietoja luokitteluun käyttävien järjestelmien käyttäjien on kerrottava ihmiselle, johon niitä käytetään, että heihin käytetään ao. järjestelmää.
- Tekoälyllä luodun aitoa muistuttavan kuva- tai äänisisällön yhteydessä on kerrottava, että kyse tekoälyllä luodusta sisällöstä (ns. deep fake).
- Mitä edellä on mainittu ei sovelleta, jos tarkoituksena on rikosten paljastaminen, estäminen tai selvittäminen.

Korkean riskin tekoälyjärjestelmiin kohdistuvat keskeiset vaatimukset



9 artikla
Riskinhallintajärjestelmä

- Suuririskisille tekoälyjärjestelmille on perustettava, pantava täytäntöön ja dokumentoitava riskinhallintajärjestelmä ja sitä on pidettävä yllä.
- Riskinhallintajärjestelmän on koostuttava iteratiivisesta prosessista, joka jatkuu suuririskisen tekoälyjärjestelmän koko elinkaaren ajan ja joka on saatettava säännöllisesti ja järjestelmällisesti ajan tasalle.
- Suuririskiset tekoälyjärjestelmät on testattava tarkoituksenmukaisimpien riskinhallintatoimenpiteiden määrittämiseksi. Testauksella on varmistettava, että suuririskiset tekoälyjärjestelmät toimivat johdonmukaisesti käyttötarkoituksensa mukaisesti ja että ne ovat tässä luvussa vahvistettujen vaatimusten mukaisia.

10 artikla
Data ja datahallinto

- Koulutus-, validointi- ja testausdatajoukkoihin on sovellettava asianmukaisia datahallinto- ja hallintakäytäntöjä.
- Koulutus-, validointi- ja testausdatajoukkojen on oltava merkityksellisiä, edustavia, virheettömiä ja täydellisiä.
- Sellaisten suuririskisten tekoälyjärjestelmien kehittämiseen, joissa ei hyödynnetä malleja kouluttavia tekniikoita, on sovellettava asianmukaisia datahallinto- ja hallintakäytäntöjä.

11 artikla
Tekninen dokumentaatio

- Suuririskisen tekoälyjärjestelmän tekninen dokumentaatio on laadittava ennen kuin järjestelmä saatetaan markkinoille tai otetaan käyttöön, ja se on pidettävä ajan tasalla.
- Tekninen dokumentaatio on laadittava siten, että siinä osoitetaan suuririskisen tekoälyjärjestelmän olevan tässä luvussa vahvistettujen vaatimusten mukainen ja annetaan kansallisille toimivaltaisille viranomaisille ja ilmoitetuille laitoksille kaikki tarvittavat tiedot sen arvioimiseksi, täyttääkö tekoälyjärjestelmä kyseiset vaatimukset.

12 artikla
Tietojen säilyttäminen

- Suuririskiset tekoälyjärjestelmät on suunniteltava ja kehitettävä siten, että ne mahdollistavat tapahtumien automaattisen tallentamisen (’lokitiedot’), kun suuririskiset tekoälyjärjestelmät ovat toiminnassa. Näiden lokitusvalmiuksien on oltava tunnustettujen standardien tai yhteisten eritelmien mukaisia.
- Lokitusvalmiuksilla on varmistettava sellainen tekoälyjärjestelmän toiminnan jäljitettävyyden taso koko järjestelmän elinkaaren ajan, joka on oikeassa suhteessa järjestelmän käyttötarkoitukseen.

13 artikla
Läpinäkyvyys ja tietojen antaminen käyttäjille

- Suuririskiset tekoälyjärjestelmät on suunniteltava ja kehitettävä siten, että varmistetaan, että niiden toiminta on riittävän läpinäkyvää, jotta käyttäjät voivat tulkita järjestelmän tuloksia ja käyttää niitä asianmukaisesti.
- Suuririskisten tekoälyjärjestelmien mukana on oltava käyttöohjeet asianmukaisessa digitaalisessa tai muussa muodossa, ja niissä on annettava ytimekkäitä, täydellisiä, paikkansapitäviä ja selkeitä tietoja, jotka ovat käyttäjien kannalta olennaisia, esteettömiä ja ymmärrettäviä.

14 artikla
Ihmisen suorittama valvonta

- Suuririskiset tekoälyjärjestelmät on suunniteltava ja kehitettävä siten, että luonnolliset henkilöt voivat tehokkaasti valvoa niitä tekoälyjärjestelmän käytön aikana, mukaan lukien asianmukaisten käyttöliittymävälineiden käyttö.
- Ihmisen suorittamalla valvonnalla on pyrittävä ehkäisemään tai minimoimaan terveydelle, turvallisuudelle tai perusoikeuksille aiheutuvat riskit, joita voi syntyä, kun suuririskistä tekoälyjärjestelmää käytetään käyttötarkoituksensa mukaisesti tai kohtuudella ennakoitavissa olevissa väärinkäyttöolosuhteissa, erityisesti jos tällaiset riskit jatkuvat huolimatta muiden tässä luvussa vahvistettujen vaatimusten soveltamisesta.

15 artikla
Tarkkuus, luotettavuus ja kyberturvallisuus

- Suuririskiset tekoälyjärjestelmät on suunniteltava ja kehitettävä siten, että ne saavuttavat käyttötarkoitukseensa nähden asianmukaisen tarkkuuden, luotettavuuden ja kyberturvallisuuden tason ja toimivat tässä suhteessa johdonmukaisesti koko elinkaarensa ajan.
- Suuririskisten tekoälyjärjestelmien on siedettävä virheitä, vikoja tai epäjohdonmukaisuuksia, joita saattaa esiintyä järjestelmässä tai ympäristössä, jossa järjestelmä toimii, erityisesti siksi, että järjestelmät ovat vuorovaikutuksessa luonnollisten henkilöiden tai muiden järjestelmien kanssa.
- Suuririskisten tekoälyjärjestelmien luotettavuus voidaan saavuttaa teknisillä vararatkaisuilla, joihin voivat kuulua varasuunnitelmat tai vikavarmistussuunnitelmat.

17 artikla
Laadunhallintajärjestelmä

- Suuririskisten tekoälyjärjestelmien tarjoajien on otettava käyttöön laadunhallintajärjestelmä, jolla varmistetaan tämän asetuksen noudattaminen. Järjestelmä on dokumentoitava järjestelmällisesti ja täsmällisesti kirjallisiksi periaatteiksi, menettelyiksi ja ohjeiksi.

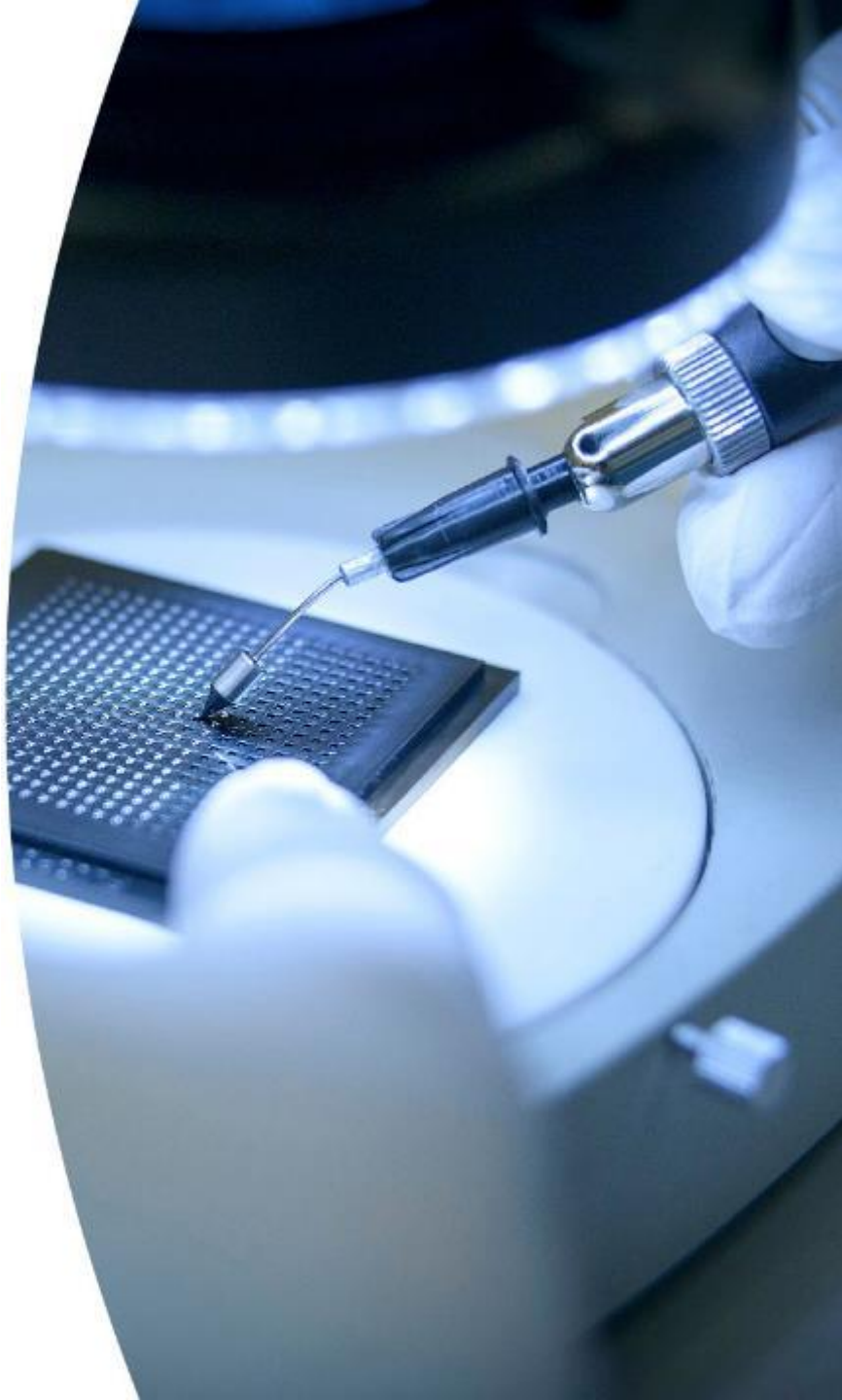
21 artikla
Korjaavat toimenpiteet

- Suuririskisten tekoälyjärjestelmien tarjoajien, jotka katsovat tai joilla on syytä uskoa, että niiden markkinoille saattama tai käyttöön ottama järjestelmä ei ole tämän asetuksen mukainen, on välittömästi toteutettava tarvittavat korjaavat toimenpiteet kyseisen järjestelmän saattamiseksi vaatimusten mukaiseksi, sen poistamiseksi markkinoilta tai sitä koskevan palautusmenettelyn järjestämiseksi.

29 artikla
Suuririskisten tekoälyjärjestelmien käyttäjien velvollisuudet

- Suuririskisten tekoälyjärjestelmien käyttäjien on käytettävä tällaisia järjestelmiä niiden mukana seuraavien käyttöohjeiden mukaisesti.
- Siltä osin kuin käyttäjä valvoo syöttötietoja, käyttäjän on varmistettava, että syöttötiedot ovat merkityksellisiä suuririskisen tekoälyjärjestelmän käyttötarkoituksen kannalta.
- Käyttäjien on seurattava suuririskisen tekoälyjärjestelmän toimintaa käyttöohjeiden perusteella.
- Suuririskisten tekoälyjärjestelmien käyttäjien on säilytettävä suuririskisen tekoälyjärjestelmän automaattisesti tuottamat lokitiedot siltä osin kuin tällaiset lokitiedot ovat niiden määräysvallassa.

Tekoälyasetuksen ja kansallisen sääntelyehdotuksen vertailua




Kielletyt käyttötarkoitukset (Art 5)
Ei juuri hallinnon APT:n käyttöä
Korkean riskin käyttötarkoitukset (korkean riskin järjestelmät) (Art 8–50)
Osa hallinnon APT:n käyttötarkoituksista
Käyttötarkoitukset, joihin kohdistuu läpinäkyvyysvelvoite (Art 52)
Hyvin pieni osa hallinnon APT:n käyttötarkoituksista
Kaikki muut käyttötarkoitukset (”matalan riskin” järjestelmät) – ei sääntelyn piirissä
Osa hallinnon APT:n käyttötarkoituksista

Ehdotus tekoälyasetukseksi	Kansallinen ehdotus APT-sääntelyksi
Riskinhallintajärjestelmä (art 9)	TiHL 28 b ja 28 c § (osittain)
Datahallinto (art 10)	TiHL 28 h § ja voimassa oleva TiHL 4 luku
Tekninen dokumentaatio (art 11)	TiHL 28 a § ja 28 d §
Tietojen säilyttäminen (art 12)	TiHL 28 c § ja voimassa oleva 17 §
Läpinäkyvyys ja tietojen antaminen käyttäjille (art 13)	*Ei vastaavaa kansallista sääntelyä (kyse läpinäkyvyydestä käyttäjälle, ei käyttäjän asiakkaalle)*
Ihmisen suorittama valvonta (art 14)	TiHL 28 c §
Tarkkuus, luotettavuus ja kyberturvallisuus (art 15)	TiHL 28 b § ja 28 c § (ainakin osin), voimassa oleva sääntely (esim. TiHL 4 luku)
Laadunhallintajärjestelmä (art 17)	TiHL 28 a §, 28 c §, 28 d §
Korjaavat toimenpiteet (art 21)	*Ei suoraan vastaavaa sääntelyä*
Suuririskisten tekoälyjärjestelmien käyttäjien velvollisuudet (art 29)	Koko 6 a luku (tavallaan)

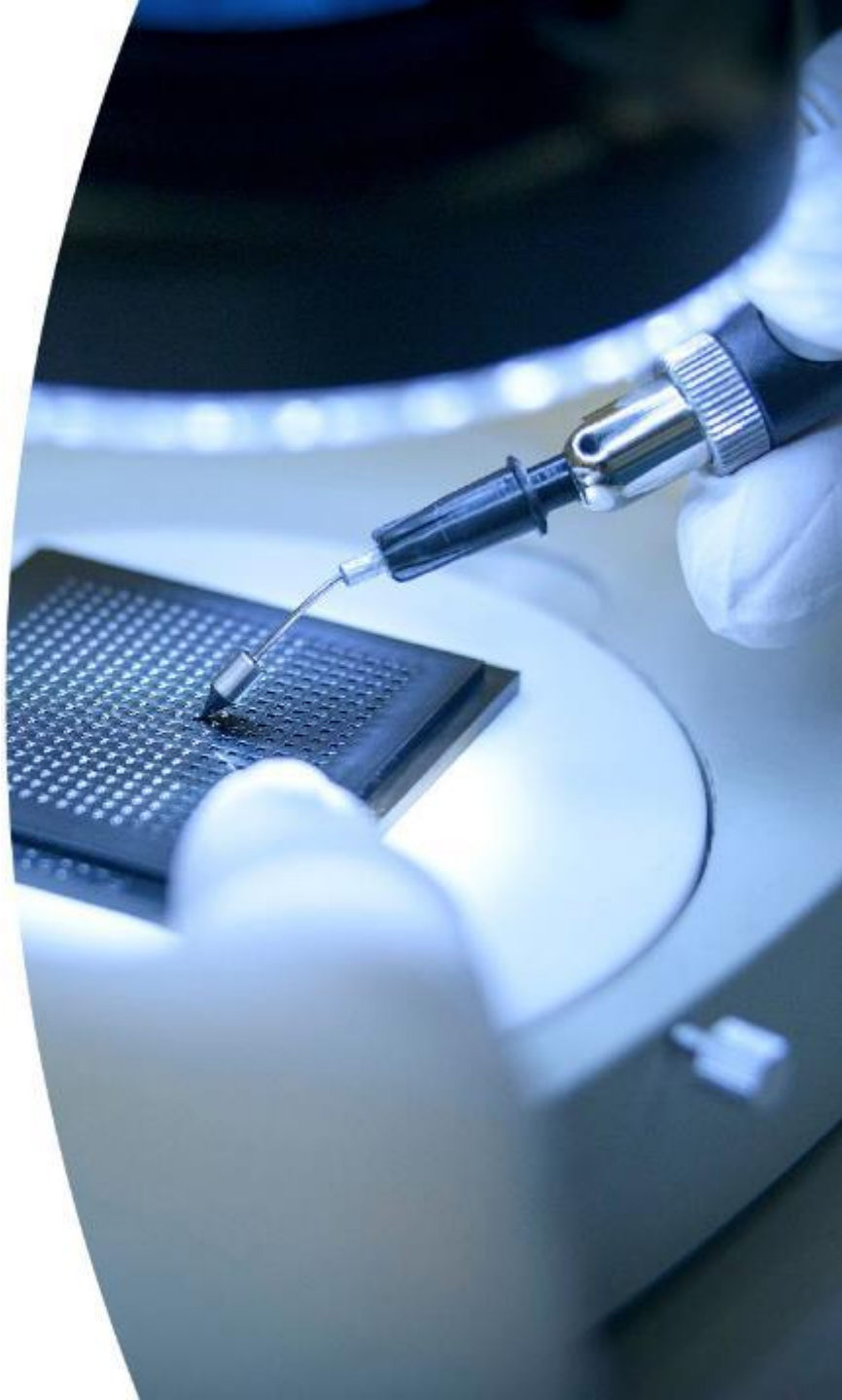
Keskeisiä eettisiä periaatteita tässä asiayhteydessä (lähde: VM julkaisuja 54:2021)

- Luotettavuus
- Ymmärrettävyys ja selitettävyys
- Asianmukaisuus ja tarkoituksenmukaisuus
- Yhdenvertaisuus
- Yksityisyydensuoja
- Itsemääräämisoikeus
- Turvallisuus ja koskemattomuus
- Bonus: virkavastuun kohdistuminen perustuslain mukaisesti

Miten eettisten periaatteiden toteutumista edistetään sääntelyehdotuksissa?

Eettinen periaate	Tekoälyasetus edistää?	Kansallinen uusi sääntely edistää?
Luotettavuus	Kyllä	Kyllä
Ymmärrettävyys ja selitettävyys	Ei	Kyllä
Asianmukaisuus ja tarkoituksenmukaisuus	Ehkä?	Kyllä
Yhdenvertaisuus	Ehkä?	Kyllä
Yksityisyydensuoja	Ei	Ei
Itsemääräämisoikeus	Ei	Ehkä?
Turvallisuus ja koskemattomuus	Kyllä	Ehkä?
Bonus: virkavastuun kohdentuminen	Ei	Kyllä

Ajatuksia



Ajatuksia 1/3

- Tekoälyn määritelmä ja siten asetuksen soveltamisala on ongelmallisen laaja ja epäselvä. Jää nähtäväksi, millainen siitä muotoutuu valmistelun aikana.
- Vaikka asetusta sovellettaisiinkin kaikkiin monimutkaisempiin IT-järjestelmiin, ei sen sääntely siltikään vaikuttaisi huomattavaan osaan hallinnon APT-käyttötarkoituksista (esim. verotukseen tai kaupparekisteriin), koska ne eivät ole ”korkeariskisiä”.

Ajatuksia 2/3

- Tekoälyasetus rakentuu vahvasti riskiperustaisuudelle. Vaikka tässä ei olekaan juuri hienosyisyyttä, on siinä silti enemmän sävyeroja kuin kansallisessa sääntelyssä.
- Tekoälyasetus ei tarjoa ratkaisuja vastuun (virkavastuun) kohdistumisen, hyvän hallinnon periaatteiden varmistamisen tai julkisuusperiaatteen ongelmiin. Tämä johtunee yhtäältä asetuksen perimmäisestä luonteesta tuoteturvallisuussääntelynä ja toisaalta siitä, ettei tämä ole edes EU-sääntelyn tarkoitus.

Ajatuksia 3/3

- Tekoälyn eettisistä periaatteista on puhuttu ja julkaistu paljon, mutta ne esiintyvät tekoälyasetuksessa lopulta varsin vähäisessä roolissa.
  - Ihmiskeskeisyyttä, kansalaisten oikeuksien toteutumista yms. asetuksessa edistetään erityisen vähän. Sääntelyn tarkoituksena on ennemmin selkeyttää tekoälyjärjestelmien tuottajien ja käyttäjien välisiä (valta)suhteita.
- Toisin kuin tekoälyasetus, kansallinen sääntely kohdistuisi teknologian ja tuotteiden sijasta hallintomenettelyyn ja viranomaisen toimintaan. Sitä kautta se on voitu rakentaa eettisiä periaatteita ja perusoikeuksia jo nyt toteuttavien kansallisen sääntelyn kerrosten päälle (esim. perustuslaki ja yhdenvertaisuuslaki).

VALTIOVARAINMINISTERIÖ
FINANSMINISTERIET

fin