

A large, abstract green graphic that resembles a stylized map of Finland or a similar geographical shape, positioned at the top of the slide. It has a jagged, irregular bottom edge.

AuroraAI-esiselvityshanke

Etiikka-työkokonaisuuden suositukset

Meeri Haataja – Saidot

Marko Latvanen – VRK

AuroraAI:n Etiikka-verkosto

#AURORA**AI**

Sisällys

1. Työkokonaisuuden tavoitteet
2. Työpaketeissa tunnistettuja eettisiä avainkysymyksiä
3. Alustava eettinen koodisto
4. Lainsäädännöllinen perusta
5. Suositukset eettisen koodiston toimeenpanemiseksi

Työkokonaisuuden tavoitteet

- Tunnistaa AuroraAI:n keskeiset eettiset kipukohtat elämäntapapiloteista saatujien kokemuksien pohjalta
- Määrittää verkostotyönä AuroraAI:lle alustava eettinen koodisto pohjaksi alueen jatkotyölle
- Tunnistaa Aurora-palvelumallin eettisen koodiston toimeenpanoa koskevat keskeiset kehityskohteet

Auroran työpaketeissa ja työpajoissa tunnistettuja eettisiä kipukohtia

Yksityisyydensuoja: Kuinka yksityisyydensuojan prosesseja ja toimintatapoja hallitaan Aurora-verkostossa? Tulisiko toimintamallien ja prosessien olla verkostolle yhteisiä? Kuinka organisaatioiden vaihteleva tietosuojamaturiteetti hallitaan?

Vastuut: Kuka vastaa Aurora-verkoston tuottamista suosituksista tai päätöksistä? Tiedolle on oltava määritetty rekisterinpitäjä, joka vastaa tiedosta ja sen oikeellisuudesta sekä tekoälyn tekemistä suosituksista/päätöksistä. Rekisterinpitäjien, yhteisrekisterinpitäjien ja käsittelijöiden välisen vastuunjaon kirjaaminen kriittistä. Sopimukset ekosysteemeihin liittymisille (ml. oikeudet ja vaatimukset).

Suostumusten hallinta verkostossa: Viranomaistiedon luvittamiseen liittyvät tietosuojajuridiset kysymykset. Kuinka hallinnoimme suostumuksia ja varmistamme tiedon käyttötarkoituserusteisuuden toteutumisen palveluita tuottaessa julkisen ja yksityisen sektorin toimijoiden verkostona?

Datan luotettavuus ja laatu: Kuinka varmistamme käytettävän datan luotettavuuden ja sen käyttötarkoituksen vaatiman laadun ja ajantasaisuuden? Tarvitsemme tiedon laadun kriteerit Aurora-verkkoon liittymiselle? Kuinka tieto yhdenmukaistetaan?

Suosittelujen ja päätösten laatu: Mitkä kysymykset soveltuvat Auroran suositusten piiriin, ja kuinka varmistamme ja valvomme suositusten ja päätösten laatua? Vaikutukset saattavat näkyä usein vasta vuosien päässä tulevaisuudessa.

Läpinäkyvyys, selitettävyys, ymmärrettävyys: Mitä tietoja kansalaisille tulisi tarjota Auroran palveluista ja niiden taustalla olevasta teknologiasta, ja kuinka tämä tehdään ymmärrettävästi? Auroran suositusten ja valintojen perustelut, sekä palveluiden kehittäjät ja operoijat oltava läpinäkyviä. Kansalaisen tiedettävä asioivansa sovelluksen eikä ihmisen kanssa.

Yhdenvertaisuus, tasavertaisuus ja erityiset kohderyhmät: Onko Aurora tarkoitettu kaikille vai vain osalle kansalaisista? Kuinka yhdenvertaisuus ja erityiset käyttäjäryhmät, mm. lapset, vammaiset, maahanmuuttajat, huomioidaan palveluiden saatavuudessa? Digi-esteettömyyden vaatimus, saavutettavuus, missä vaiheessa? Miten ratkaistaan laitteisiin, osaamisiin, yhteyksiin ja tukeen liittyvät tarpeet. Kehittävien tiimien monimuotoisuus ja kansalaisten osallistaminen.

Kansalaisten osaaminen: Mitä uutta osaamista Auroran palveluiden käyttäminen vaatii kansalaisilta, voidaanko ihmisiltä odottaa tarvittavan osaamisen omaksumista? Kuinka osaamisen kasvattamista tuetaan? Miten oma dataprofiini voi vaikuttaa minuun? Miten olen "itseni tiedolla johtaja"?

Hallinnon ja kansalaisen suhde: Kuinka Aurora muuttaa kansalaisen ja hallinnon suhdetta? Minun hyväni ja tavoitteeni vs. yhteiskunnan hyvä ja tarpeet - miten nämä sovitetaan yhteen? Luottamus ja Auroran vision ja toteutuksen vaikutukset hallinnon ja kansalaisen väliseen luottamukseen. Kuinka Auroraa valvotaan?

Alustava eettinen koodisto

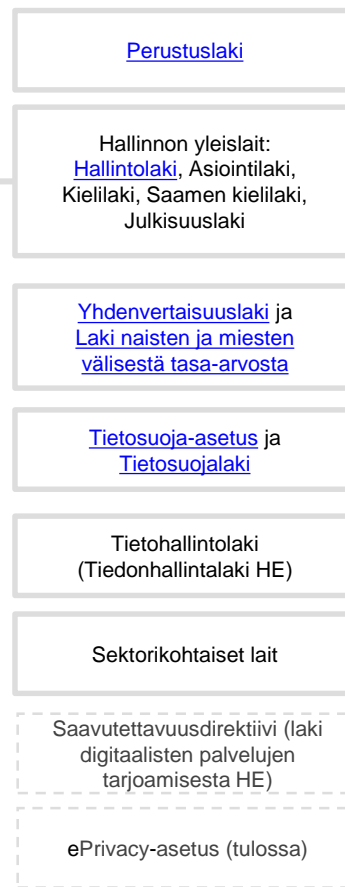
1. Ihmiskeskeinen ja ihmisarvoja kunnioittava tarkoitus
2. Yhdenvertaisuus ja syrjimättömyys
3. Yksityisyydensuoja
4. Tiedon laatu ja hallinta
5. Vikasietoisuus ja turvallisuus
6. Tiedollinen itsemääräämisoikeus
7. Omistajuus ja vastuut
8. Läpinäkyvyys

Lain- säädn- nöllinen perusta

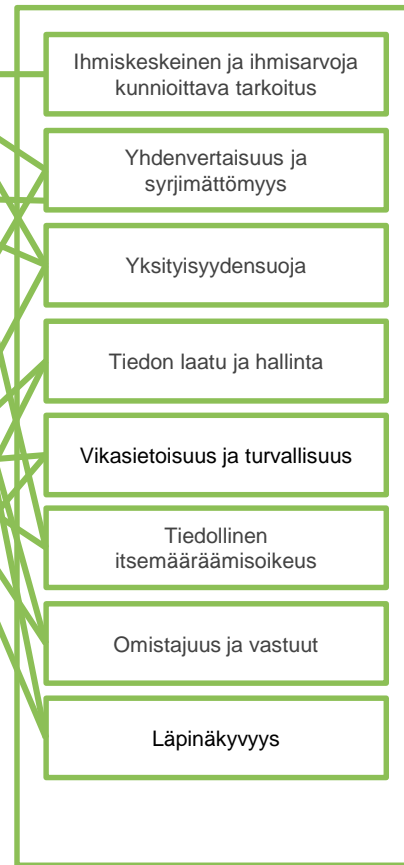
Hyvän hallinnon periaatteet (Hallintolaki)



Keskeinen lainsäädäntö



Eettinen koodisto



Suosituksukset

AuroraAI:n toteutettavuuden
kannalta kriittiset kokonaisuudet

AuroraAI:n juridinen viitekehys

Regulatiivinen MyData-hiekkalaatikko tiedon
toissijaisen käytön mahdollistamiseksi

AuroraAI:n vastuullista toteutusta
mahdollistavat ja ohjaavat kokonaisuudet

Hallinnollinen organisoituminen,
omistajuudet, vastuut ja valvonta

Läpinäkyvyyden vaatimukset
ja toteutumisen käytännöt

Eettiset suunnittelukäytännöt ja vaikutusarvio

Kansalaisten osallistuminen ja osaaminen

Aurora-verkoston monimuotoisuus
ja etiikkaosaamisen kasvattaminen

Tekniset välineet eettisyyden varmistamiseen
AuroraAI-arkkitehtuurissa

Etiikka-
vetovastuu ja
etiikkaverkosto

Aurora AI:n juridinen viitekehys

Juridisen viitekeh്യksen tehtävä on ohjata ja tukea AuroraAI-verkoston toiminnan lainmukaisuutta ja luoda mekanismit juridisten kysymysten tunnistamiselle, käsittelylle ja edistämiseksi yli hallinonalojen.

Juridisesta viitekehuksesta vastaa esimerkiksi OM:n ohjauksessa toimiva, hallinnonalat ylittävä juristiverkosto.

Verkostossa on juridista asiantuntemusta kaikista AuroraAI:n olennaisista osa-alueista, joista kategorisesti edellytään teknologia, sopimusjuridiikka, rekisterinpito ja tietoluovutukset, tietosuoja, ihmis- ja perusoikeudet ml. tasa-arvo ja yhdenvertaisuus.

Regulatiivinen MyData-hiekkalaatikko tiedon toissijaisen käytön mahdollistamiseksi

On tunnistettu tarve luoda regulatiivinen MyData-hiekkalaatikkoympäristö julkisen hallinnon keräämän henkilötiedon toissijaisen käytön mahdollisuuksien tutkimiseen rajatussa ympäristössä, sekä valmistelemaan tällaisten uudenlaisten tiedon käyttötärpeiden mahdollistamisen vaatimat lainsäädäntömuutokset poikkihallinnollisesti.

Esitetyn hiekkalaatikkoympäristön tavoite on

1. Mahdollistaa julkisen sektorin MyData-kokeilujen toteuttaminen rajatussa ympäristössä kansalaisten vapaaehtoisuuteen ja suostumuksiin perustuen
2. Tutkia, tuoko uusi teknologia sellaisia uusia kansalaisille hyödyllisiä mahdollisuuksia julkisen sektorin keräämän datan hyödyntämiselle, joiden merkittävyyden vuoksi on tarve laajentaa viranomaisten lakisääteisiä velvoitteita tai mahdollistaa näiden ulkopuolinen datan toissijainen hyödyntäminen esimerkiksi kansalaisen suostumuksella
3. Valmistella mahdollisten lainsäädäntömuutosten ketterä toteuttaminen poikkihallinnollisesti

Hallinnollinen organisoituminen, omistajuudet, vastuut ja valvonta

AuroraAI:lla on oltava yksiselitteinen omistajaorganisaatio, joka on viimekätisessä vastuussa hallintamallista (ml. toimijaroolien vastuut), kokonaisuuden lainmukaisesta toiminnasta ja eettisen koodiston toteutumisesta.

Omistaja voi asettaa AuroraAI:lle ohjausryhmän ja tahoja, jotka osaltaan vastaavat näiden vaatimusten toteutumisesta. Tämän suosituskoodiston tekijät esittävät dedikoidun, monialaisen asiantuntijaryhmän asettamista eettiseksi seurantaryhmäksi AuroraAI:lle.

Aurora-verkossa toimivilta palveluntuottajilta vaaditaan sitoumus eettisen koodiston noudattamisesta, ja sitoumusrikkomuksiin on asetettava menettelyt ja sanktiot. Tässä, kuten vastuukysymysten määrittelyssä muutenkin, tarvitaan suosituksessa 1. mainitun juristiverkoston osallistumista.

Läpinäkyvyyden vaatimukset ja toteutumisen käytännöt

Kansalaisen kannalta ehdottomia vaatimuksia läpinäkyvyydelle ovat omistaja- ja kontaktitahojen tuominen selvästi esille, palvelun toimintalogiikan selittäminen mahdollisimman yleistajuisesti, palveluntuottajaverkossa olevien toimijoiden tarjoaminen listana tai tiedostona sekä täsmällinen ohje palvelun käyttöön ja käyttäjän oikeuksiin virhe- tai häiriötilanteissa.

Algoritmien toimintalogiikkaa koskeva läpinäkyvyys on haaste: varsinkaan yksityisten palveluntuottajien algoritmeja ei yleensä päästetä toisten nähtäviksi. Muutenkin algoritmeihin perustuvan koneellisen päätöksenteon perusteet voivat olla erittäin vaikeasti selvitettävissä. Tämä osa-alue vaatii erityistä analysointia, koska emme voi hyväksyä tilannetta, jossa joudumme sokkona hyväksymään koneen tuottamaa neuvontaa tai päätöksiä.

Eettiset suunnittelukäytännöt ja vaikutusarvio

Eettisillä suunnittelukäytännöillä tarkoitetaan käytäntöjä, joilla eettiset periaatteet realisoidaan vaatimuksiksi ja toteutusta ohjaaviksi päätöskriteereiksi tekoälypalveluiden kehitys- ja käyttöprosesseissa. Konkreettinen esimerkki on eettinen vaikutusarviointi, jollaisia on alettu hyödyntää erityisesti julkisen sektorin tekoälykehityksessä mm. Kanadassa.

On suositeltavaa, että AuroraAI:lle määritetään koko verkostoa sitovat toimintatavat tekoälysovellusten eettiseen vaikutusarviointiin, ja näiden toteuttaminen vastuutetaan selkeästi.

Tekoälyjärjestelmän eettisessä suunnittelussa on huomioitava koko kaari konseptista loppukäyttäjäpalvelun muotoiluun. Eettisyyden integrointi systeemiin on merkittävä haaste; meillä on silti oltava eettinen ohjeisto ja sitä on käytettävä eettisen auditoinnin pohjana. Vaikutusarviossa on huomioitava useat osa-alueet: vaikutus yksilöihin, heijastusvaikutukset lainsäädäntöön, ympäristöön, työmarkkinoihin, koulutustarpeeseen jne.

Eettisen suunnittelun kaaren lopussa on huomioitava yhdenvertaisuuteen ja syrjimättömyyteen liittyvät tekijät palvelumuotoilun, esteettömyyden ja käytettävyyden osalta: vähemmistöryhmät, ikääntyneet, kielet, erilaiset päätelaitteet jne. Samalla on huomioitava ICT-infrastruktuuri Suomessa.

Kansalaisten osallistaminen

Koska kansalaisyhteiskunnassa ei ole vielä ”konsensusymmärrystä” tekoälystä ja sen ominaisuuksista, on välttämätöntä aktiivisesti paitsi kouluttaa, myös osallistaa kansalaisia, ml. potentiaalisen syrjinnän tai tahattoman ulos sulkemisen riskissä olevat väestöryhmät.

Laajat kansalaiskuulemiset ovat välttämättömiä, liittyen kaikkiin AuroraAI:n osa-alueisiin: palvelun tavoitteet ja toimintaperiaatteet, palvelumuotoilu jne.

Menetelminä on oltava koko osallistavan suunnittelun työkalupakki: kyselyt, fokusryhmät, työpajat ja tapaamiset, osallistamissivustot ja -palvelut. Kertyvä aineisto on dokumentoitava ja raportoiva ja sen soveltamisesta kehittämistyöhön on viestittävä avoimesti.

Aurora-verkoston monimuotoisuus ja etiikkaosaamisen kasvattaminen

Aurorassa erilaiset järjestelmätoimittajat ja palvelutuottajat kytkeytyvät monenkeskiseen tekoälyjen verkkoon. Eettisinä haasteina ovat erilaisten toimijoiden motiivien yhteensovittaminen sekä eettisten valmiuksien ja asiassa sisäistetyn tietomäärän erot.

Suosittelaa ohjemateriaaleja ja koulutuksia eettisen AI:n perusteista ja toteuttamisesta. Koulutuksen päätteeksi Aurora-toimija saa sertifi kaatin, jolla se todistaa olevansa tekoälyssä vastuullinen. Sertifi kaatin saaminen olisi julkinen ja pitävä sitoutuminen noudatettaviin periaatteisiin.

Kansalaisten tietoisuutta AuroraAI:n teemoista on nostettava esim. Tiedotuskampanjalla vapaan sivistystoiminnan ICT- ja AI-kurssien yhteydessä. On löydettävä menettelyjä, joiden avulla kansalaisyhteiskunta ei jää passiiviseksi, kyvyttömäksi vastaanottajaksi AI-palvelujen vahvistuessa.

Tekniset välineet eettisyyden varmistamiseen AuroraAI-arkkitehtuurissa

Jo tällä hetkellä saatavilla on mm. apuvälineitä datan ja algoritmien vinoumien testaamiseen ja selitettävyyteen. AuroraAI:ssa on varmistettava, että arkkitehtuurin kehittämisessä tutkitaan tällaisten teknisten välineiden tarjoamia mahdollisuuksia, ja otetaan käyttöön sopivia teknisiä välineitä tukemaan eettisen koodiston käytännön toteuttamista Aurora-arkkitehtuurissa.

On suositeltavaa myös jatkaa jo käynnistettyä työtä AuroraAI:n kytkemisessä kansainväliseen tekoälyn etiikan standardointi- ja sertifiointityöhön, ja edistää tätä kautta Aurora-arkkitehtuurin lähtökohtaista eettistä auditoitavuutta ja sertifioidtavuutta. On huomioitavaa, että AuroraAI voi tarjota myös merkittävän ympäristön edistää suomalaisia innovaatioita ja tutkimusta tekoälyn etiikan tutkimuksen ja kehityksen alueella.

Kiitos

AuroraAI:n Etiikka-verkoston osallistuneet
asiantuntijat elämäntapapiloteista

Ministeriöiden ja virastojen tietosuojajuristit ja muut
asiantuntijat

Yhdenvertaisuusvaltuutetun toimiston henkilöstö

FCAI Society –verkoston jäsenet

Suomi.fi kansalaispaneelin kyselyyn osallistuneet
kansalaiset

AuroraAI-Slackissa teeman keskusteluihin
osallistuneet

#AURORA**AI**