DIGI- JA VÄESTÖTIETO-VIRASTO

Tekoälyn hyödyntäminen – huoneentaulut ja tarkistus-listat

VAHTI hyvät käytännöt -tukimateriaali

13.6.2024 versio 2.5

Sisällysluettelo

Tekoälyn hyödyntäminen – huoneentaulut ja tarkistuslistat

1. Johdanto

Tämä tukimateriaali on laadittu julkisen hallinnon organisaatioille turvallisen työskentelyn ja toiminnan edistämiseksi. VAHTI hyvät käytännöt -tukimateriaalit pohjautuvat Julkisen hallinnon digitaalisen turvallisuuden johtoryhmän (VAHTI, https://dvv.fi/vahti) asiantuntijaryhmien kokoamiin suosituksiin ja hyviin käytäntöihin riskienhallinnan, toiminnan jatkuvuuden ja varautumisen, osaamisen kehittämisen, ICT-palveluiden, tietoturvallisuuden, sekä tietosuojan alueilla. Niiden mukaan toimimalla edistämme samalla kyberturvallisuuden toteutumista.

VAHTI hyvät käytännöt -tukimateriaalit ovat ensisijaisesti suunnattu julkisen hallinnon organisaatioille, mutta ne ovat vapaasti minkä tahansa organisaation hyödynnettävissä. Toivomme, että mikäli kehitätte tai parannatte näitä materiaaleja, annatte niistä myös palautetta sisältöjen edelleen kehittämiseksi. Otamme mielellämme vastaan parannus- ja korjausehdotuksia ja julkaisemme päivitetyn version, kun niitä on kertynyt riittävästi. Voit lähettää palautetta: digiturva@dvv.fi – kirjoita otsikkoon ”VHK tekoäly huoneentaulut”.

Tämän dokumentin ensimmäinen versio muodostettiin hyödyntämällä muun muassa tekoälyä ja se julkaistiin 15.6.2023. Ensimmäinen asiantuntijoiden jatkokehittämä versio (1.0) julkaistiin 12.9.2023. Version 2.0 kehittämiseen osallistui 17.11.2023 työpajassa 100 asiantuntijaa VAHTI-työryhmistä ja DVV:ltä, ja se julkaistiin 12.12.2023. Versiossa 2.0 on tehty päivityksiä listojen rakenteeseen ja listatut kohdat on numeroitu viittausten helpottamiseksi (kyse ei ole tärkeysjärjestyksestä tai vastaavasta). Huoneentauluista on myös tuotettu työpajassa toivotun mukaisesti tiivistetty diasarja.

13.6.2024 julkaistu versio 2.5. pohjautuu valmistelutyöryhmän ja kaikille avoimen ennakkokyselyn pohjalta päivitettyyn versioon.

VAHTI hyvät käytännöt tukimateriaalien tuottamisessa hyödynnämme apunamme tekoälytyökaluja, kuten on aiemmissa versioissa. Kaikki sisällöt käyvät läpi tarkistus- ja kommentointiprosessin tehty tämän dokumentin.

Julkaisun jakelu: https://dvv.fi/digiturvajulkaisut

2. Huoneentaulut ja tarkistuslistat

Organisaatiot voivat vapaasti hyödyntää, sovittaa ja muokata omiin tarpeisiin tässä materiaalissa olevia huoneentauluja, jotka on kohdistettu eri toimijaryhmille suunnattuun ohjaukseen, viestintään ja koulutukseen. Ne toimivat tarkistuslistoina keskeisiksi tunnistetuista asioista, mutta ne eivät välttämättä kata kaikkia näkökulmia, joita eri organisaatioissa voi tulla vastaan.

Organisaatioiden tulee ohjeistaa tarkemmin käyttöön hyväksytyistä tekoälypalveluista ja niiden turvallisesta käytöstä. Jokainen organisaatio ja asiantuntija vastaa siitä, että tämän tukimateriaalin sisältö sovitetaan vastaamaan organisaation omaa toimialaa ja sitä koskevaa lainsäädäntöä. Materiaalia ei saa ottaa käyttöön sellaisenaan ilman läpikäyntiä, muokkaamista omaan tarkoitukseen sekä siihen liittyvää asianmukaista viestintää organisaatiossa ja tarvittaessa sen sidosryhmille.

Tekoälylle ei ole yhtä selkeää määritelmää, alla neljä sen toimintaan liittyvää keskeistä havaintoa:

- Tekoäly on tiettyä käyttöä varten laadittu **tietokoneohjelma.**
- Se **vaatii toimiakseen tietokoneen** eli muun muassa prosessorin tai valtavan määrän prosessoreita tai muita GPU-laskentayksiköitä, ram- ja apumuisteja (levyjärjestelmä) tietojen käsittelemiseksi ja tallentamiseksi, suorituskykyiset tietoliikenneyhteydet ja muuta teknologiaa. Siis bittejä, ”rautaa” ja sähköä.
- **Tekoälyllä ei ole tunteita**, **tietoisuutta ja itsetietoisuutta**, mutta se voi näytellä pyydettäessä niitä ja muutenkin se voi pyrkiä pyydettäessä toimimaan kuten ihmiset, myös virheellisesti. Tekoäly on eräänlainen näyttelijä, tarvittaessa.
- Kaikki sen tuottama tieto – kuvat – musiikki – videot – puhe pohjautuvat meidän ihmisten tuottamaan **opetusdataan** – josta se luo vastaukset **matemaattisten mallien, todennäköisyyksien perusteella** – hieman kuten ennustava tekstinsyöttö älylaitteissa.

Tässä materiaalissa tekoälyllä tarkoitetaan laajasti järjestelmiä, prosesseja ja palveluita, joissa hyödynnetään tekoälyä sekä joissain tapauksissa myös sellaisia kehittyneitä järjestelmiä, joiden algoritmit tai rakenne ovat siinä määrin kompleksisia, että ne ovat verrannollisia tekoälyyn tai voivat näyttäytyä sellaisina käyttäjälle.

Tällä tavallista laajemmalla rajauksella halutaan sisällyttää eri määritelmiä kattavasti sekä kiinnittää huomiota siihen, että tekoälyyn liittyvä teknologiat kehittyvät nopeasti, niiden vaikutukset ulottuvat kauan ja eri tasoille. Ei siis ole olemassa vain yhtä ”keskiarvoista” tekoälyä.

Eri käyttötapauksiin voidaan nähdä liittyvän erilaisia toiminnallisuuksia, uhkia, hallintorakenteita, sääntelyä ja odotuksia – muun muassa. Näiden eri näkökulmien kautta muodostuu erilaisia tapoja luokitella tekoälyjä, jotka ovat tapoja viestiä niihin liittyvistä asioista. On organisaatiokohtaista, mitkä näistä valikoituvat käyttöön ja miten niistä ohjeistetaan.

Organisaation tekoälynkäytön tulisi olla suunnitelmallista, ohjattua ja johdonmukaista. Tekoäly nähdään yleisesti jo niin merkittävänä ICT-teknologiana, että se tulee vaikuttamaan organisaatiotasoisiin tavoitteisiin ja toimintakykyyn. Täten ohjaavana pohjana organisaatioilla tulisi olla suunnitelma tekoälyä varten, riippumatta siitä, onko organisaatio itse eturintamassa sen käyttöönotossa vai vasta viimeisten joukossa. Tekoälyä tullaan käyttämään sen ympärillä, olemassa olevia työkaluja ja järjestelmiä tullaan kyllästämään sillä ja tekoälyä tullaan käyttämään sitä vastaan. Tekoäly on jollain aikavälillä tulossa osaksi kaikkien organisaatioiden normaalia toimintaa, mikä edellyttää muutoksia ja sen toteuttamisen johtamista. Tätä voi osin verrata internet-verkon kautta hyödynnettävien palveluiden ja erilaisten älylaitteiden avulla tuotettavien palveluiden kehittymiseen erityisesti 2000-luvun alkupuolella.

Tähän perustuen, organisaation ohjauksen tukemiseksi tekoälyä varten on suositeltavaa muodostaa suunnitelma. On suositeltavaa, mutta organisaation päätettävissä, tehdä se strategian ja politiikan tasoisina, siten, että myös niitä kehitetään erillisinä, ennen vientiä osaksi organisaation normaalia kokonaisstrategiaa ja -politiikkaa. Tekoälystrategia ja -politiikka tulisi siis muodostaa sillä tavoitteella, että ne aikanaan (linjausten, teknologian ja osaamisen maturiteetin kehityttyä), voidaan tuoda osaksi normaalia johtamista ja toimintaa.

Tekoälykehityksen pohjaksi organisaation tulisi yhdessä, johdon ohjauksessa, tunnistaa tekoälyn mahdolliset vaikutukset sisäiseen ja ulkoiseen toimintaan sekä toimintaympäristöön, jotta näitä voidaan ennakoida. Strategian tulisi kertoa organisaatiolle muun muassa mitä tekoäly sille on, mitä organisaatio tavoittelee tekoälyn käytöllä, miten tekoäly auttaa saavuttamaan organisaation tavoitteita sekä miten edetään erillisestä tekoälyn ohjauksesta sisäistettyyn normaaliin toimintaan. Organisaation tekoälypolitiikassa tulisi kuvata muun muassa miten strategian tavoitteisiin päästään, mihin tekoälyä voidaan käyttää ja mihin ei, miten tekoälytoimia hallitaan, miten organisaatiossa käytetään tekoälyä, miten tekoälyn vaikutukset käsitellään sekä miten varaudutaan toimimaan ilman tekoälyä.

Tässä dokumentissa esitetyt listaukset tarvitsevat perustakseen näitä toimintaa ohjaavia linjauksia. Organisaation tekoälytoiminnan strategiaa ja politiikkaa tulee pitää yllä ja uudistaa teknologian ja järjestelmien kyvykkyyksien muuttuessa, lainsäädännön ja muun regulaation kehittymisen mukaan sekä tekoälyn käytöstä tunnistettujen hyötyjen ja haittojen perusteella. Strategian ja politiikan välinen ero ei aina ole selkeä ja organisaatioilla voi olla erilaisia hallintokulttuureita näiden sisältöjen määrittelyssä, minkä vuoksi käytettyä jakoa voi pitää lähinnä ohjeellisena.

2.1 Huoneentaulu organisaation johdolle

1 **Kouluttaudu ja tutustu** tekoälyn perusteisiin eri näkökulmista ja ota sekä säilytä se jatkossa johdon agendalle. Ylläpidä omaa osaamistasi.
   - Valmistele visio, suunnitelma, tiekartta, strategia ja/tai politiikka organisaation toiminnan tueksi, osallistaen toimintaan mahdollisimman paljon eri tahoja

2 Tekoälyyn liittyvien roolien, vastuiden ja resurssien **määrittely sekä budjetointi**
   - **Nimeä** tekoälyjen hallinnointiin vastaava johdon edustaja ja asiantuntija(t) sekä varmista, että organisaation sovellusten hankkimisesta, niiden kehittämisestä, riskienhallinnasta, toiminnan jatkuvuudesta, tietoturvasta ja tietosuojasta vastaavat henkilöt tuntevat organisaation kehittämiin tai hyödyntämiin tekoälypalveluihin liittyvät linjaukset ja kehittämishankkeet.
   - **Varmista**, että organisaatiolla on joko omaa tai erikseen hankittavaa **erityisasiantuntemusta** lainsäädännöllisiin sekä muihin tekoälyn ympärillä tapahtuviin uudenlaisiin tilanteisiin ja ilmiöihin. Osoita riittävä varaus tekoälykoulutukseen vastuullisille asiantuntijoille sekä käyttäjille.

3 **Varmista**, että tekoälyn hyödyntämisessä ja palveluiden kehittämisessä sovelletaan organisaation käyttämää **riskienhallintamallia** sekä tunnistetut riskit käsitellään ja otetaan hallintaan.
   - **Pysy tietoisena** riskien kehityksestä nopeasti kehittyvällä kentällä ja tarkastele niitä laajasti, mukaan lukien koko palvelu- tai prosessiketju, sisältäen kokonaisuuteen liittyvät alihankinta ja toimitusketjuverkostot, johon tekoäly sisältyy.
   - Huolehdi **yhteistyöstä** sidosryhmien kanssa ja pyri yhteisiin linjauksiin
   - **Vastuuta** sopimusten tekoälyä käsittävien osioiden vastuiden hallinta
   - **Varmista**, että organisaatiossa on mahdollisuus tutkia ja kokeilla tekoälyä eri muodoissaan turvallisesti, jotta hyötypotentiaali on mahdollista tunnistaa. Varaudu myös mahdollisesti toteutuvien riskien hallintaan.

4 **Suunnittele ja toteuta muutosjohtaminen ja -viestintä**, kun tekoälyyn liittyviä tai sisältäviä kokonaisuuksia toteutetaan käyttöön (mm. vaikutukset työntekoon, asiakkaiden palveluihin, sidosryhmiin)
   - **Varmista**, että organisaation **henkilöstö tietää** ajantasaisen strategian ja politiikan sekä tuntee henkilöstön ohjeistuksen.

2.2 Huoneentaulu tekoälyn käyttöä, tietoturvallisuutta, tietosuojaa ja muita tukitoimia ohjaaville asiantuntijoille

1 **Ohjeista,** mitä sovelluksia, järjestelmiä ja palveluita saa tai ei saa käyttää ja miksi sekä missä rajoissa käyttö on sallittua.

2 **Varmista**, että tekoälyn käyttö ja kehittäminen, ohjeistukset sekä sopimus- ja lisenssiehdot vastaavat organisaation strategiaa ja politiikkaa sekä lakia ja muita vaatimuksia

- Varmista, miten tekoäly tukee lakisääteisiä tehtäviä ja toimii sallituissa rajoissa
  - Huomioi palveluiden laajennusten, uusien moduulien tai versioiden mukana tulevat käyttöehtojen muutokset.
- Huomioi tekoälyn käytön yhteydessä mahdolliset tekijänoikeuskysymykset.
  - Huomioi syötetyn tiedon ja tiedon omistajien näkökulma
  - Huomioi tuotoksen ja siinä esiintyvän opetusdatan merkitys

3 **Arvioi riskiperustaisesti**, voiko jotain asiaa tai materiaalia käsitellä tekoälyä hyödyntäen tai tekoälyä hyödyntävässä järjestelmässä.

- Huomioi tietosuoja ja tiedon luokittelu. Ohjeista henkilöstöä sallittujen ja kiellettyjen aineistojen käytöstä. Luokitteluiden ja metatietojen tulee soveltua tekoälyllä hyödynnettäviksi
- Huomioi varautuminen tekoälyn käytön rajoituksiin, estymiseen sekä myös omatoimiseen alasajon tarpeeseen

4 **Varmista ja sovi** kirjallisesti missä määrin ja millä edellytyksillä tekoälypalveluita voidaan mahdollisesti hyödyntää hankittaessa alihankintana sovelluskehitystä tai muita palveluita.

- Varmista, ettei sovelluskehityksen yhteydessä vuodeta organisaation toiminnan kannalta kriittistä tietoa.
- Mikäli tekoälyä hyödynnetään, sovi miten tarkistus, testaus, laadunvarmistus ja kehityksen seuranta tapahtuvat.
- Selvitä tekijänoikeudellisten riskien vastuut sopimuksissa.

5 **Ohjeista ja kouluta** henkilöstöä tekoälyyn liittyvistä periaatteista, muun muassa

- Avoin ja läpinäkyvä käyttö
- Oikeudenmukaisuus, syrjimättömyys ja vinoumat
- Tietosuoja ja yksityisyys
- Turvallisuus ja luotettavuus
- Ihmiskeskeisyys
- Eettinen suunnittelu ja käyttö
- Jatkuva käyttöpalautteen ja virhetietojen käsittely sekä poikkeamatilanteet
- Tekoälyn käytöstä viestiminen
- Osallistuminen ja yhteistyö

6 **Huomioi** tekoälyn hyödyntämiseen liittyvien eettisten ohjeiden ja riskienhallinnan kehittyminen sekä päivitä organisaation ohjeistuksia ja prosesseja vastaavasti.

7 **Verkostoidu** muiden organisaatioiden ja sidosryhmien kanssa keskustellaksesi ja jakaaksesi kokemuksia, tietoa linjauksista ja saadaksesi ja antaaksesi vertaistukea.

8 **Viesti säännöllisesti** tekoälyn ajankohtaisista asioista selkeästi ja osallistavasti sisäisesti ja tarvittaessa ulkoisesti.

9 **Ohjeista**, miten tuotosten vastaanottajille tulee kertoa tekoälyn hyödyntämistä materiaalin tuottamisessa (tekstit, kuvat, videot ja kaikki muu tekoälyllä tuotettu tai sen tuottamisessa hyödynnetty materiaali).

10 Tekoälyn käytön alkuvaiheessa **arvioi tilannetta tapauskohtaisesti**. Yleisesti pätevät ohjeet muokkautuvat kokemuksen karttuessa. Oleellista on kerätä tietoa.

- Luo järjestelmä, jolla voidaan kerätä tietoa kaikista tekoälyyn liittyvistä poikkeamista. Kukin tekoälyjärjestelmä voi olla erilainen ja niissä voi olla seurattavan erityisiä ominaisuuksia, mutta huolehdi kokonaisnäkymän muodostumisesta.
- Huomioi järjestelmän yhteys ja yhdenmukaisuus tietoturvapoikkeamailmoituksiin, riskitietoihin, laadunhallintaan, asiakaspalatteisiin ja ohjelmistokehitykseen siten kuin ne ovat tehokkainta järjestää.

2.3 Huoneentaulu henkilöstölle

1 **Noudata** organisaation käyttöperiaatteita (politiikkaa, ohjeistuksia, roolituksia, prosesseja) tekoälyn kokeiluissa, käyttöönotossa ja käytössä.

- Käytä tekoälysovelluksia vain organisaation hyväksymään tarkoitukseen.

2 **Käytä** vain organisaation hyväksymiä tekoälysovellutuksia tietojen käsittelyyn.

- Mikäli käyttämäsi ohjelmisto tai verkkopalvelu ilmoittaa ottaneensa käyttöön tekoälyyn perustuvia ominaisuuksia, ilmoita tästä organisaatiossasi ohjelmistojen hyväksymisistä vastaavalle taholle.
- Tarkista, onko käyttämäsi palvelu kaikille avoin, yleinen tekoälypalvelu vai oman organisaatiosi tarjoama palvelu ja noudata niiden käytöstä annettuja ohjeita.

3 **Älä oleta, että tekoäly on aina oikeassa**. Tekoäly ei ole erehtymätön, ja sen päätökset voivat olla vääriä. Älä luota sokeasti sen päätöksiin tai tuottamaan tietoon.

- Älä käytä tekoälyn tuottamaa aineistoa (esim. koodia, dataa, käännöksiä), jos et ymmärrä mitä se tekee, sisältää ja tarkoittaa
- Tarkasta tekoälyn tuottamat tiedot muuta kautta, jos mahdollista.
- Jos et voi tai osaa itse tarkistaa tekoälyllä tuotetun tiedon tai muun materiaalin, tuotoksen oikeellisuutta, pyydä apua, älä julkaise tai hyödynnä tietoa ilman, että joku muu asiantuntija on sen oikeellisuuden varmistanut.

4 **Mikäli** tekoäly tuottaa huonolaatuista, virheellistä, vaarallista tai muuten epäilyttävää ja sopimatonta materiaalia, lopeta sen käyttö. Dokumentoi kopioimalla tekstit, ottamalla kuvakaappauksia tai tallentamalla mahdolliset virhe- ja lokitiedot analysointia varten.

- Ilmoita sovitulla tavalla tekoälyn tekemistä virheistä. Pienetkin virheet voivat kertautua huomattavaksi ongelmaksi. Organisaatiosi tulisi ohjeistaa, miten ja mistä ilmoitetaan.

5 **Varmista** aina tekoälysovellutusta tai -palvelua käyttäessäsi, ettet syötä sille organisaation rajoittamaa materiaalia.
- Esimerkkejä kielletyistä voivat olla henkilötiedot, salassa pidettävät tiedot, turvallisuusluokitellut tiedot, tekijänoikeudelliset tiedot jne.

6 **Jos olet epävarma**, kysy organisaation tekoälyvastaavalta, tietoturva-asiantuntijalta, tietosuojavastaavalta tai muulta asiantuntijalta.
- Anna palautetta tekoälyvastaavalle ja tietoturva-asiantuntijoille, jos huomaat ohjeissa ”porsaanreiän”, muun organisaatiosi luotettavaa ja turvallista toimintaa tai mainetta haittaavan ominaisuuden tai muun epäkohdan.

7 **Pidä** itseäsi ajan tasalla, osallistu aiheeseen liittyviin koulutuksiin ja muihin tilaisuuksiin.

8 **Noudata** hyvän, avoimen hallinnon periaatteita ja muista virkavastuu – vastuuta ei voi siirtää tekoälylle.
- Tekoälyn tuottamasta sisällöstä on pääsääntöisesti ilmoitettava organisaation ohjeen mukaisesti

9 Näe tekoäly positiivisena, uudenlaisena, osaamistasi laajentavana tukipalveluna.

2.4 Tarkistuslista tekoälyä sisältävän järjestelmän, palvelukokonaisuuden tai prosessin tuoteomistajille, ylläpitäjille ja kehittäjille (ym.)

TEE NÄIN:

1 **Käytä** työnantajan käyttöösi antamia tekoälypalveluita ja -resursseja annettujen ohjeiden mukaisesti työn tekemiseen ja suunniteltujen projektien toteuttamiseen.
- Kokeiluiden tekemistä tulisi tukea. Niiden on kuitenkin oltava riittävän rajattuja, riskit tulee arvioida ja riskien toteutumiseen tulee etukäteen varautua.

2 Kehitystoimien ja tekoälyn hyödyntämisen tulee **noudattaa** organisaation strategiaa ja politikkaa. Mahdolliset poikkeamat linjauksista ja kokeilut tulee käsitellä etukäteen.

3 **Noudata** lakeja ja määräyksiä. Varmista, että tekoälyn käyttö järjestelmän toiminnan apuna noudattaa kaikkia sovellettavia lakeja ja määräyksiä, mukaan lukien tietosuoja- ja syrjimättömyyslait. Päivitä tietämystäsi säännöllisesti niiden tulkinnasta ja huomioi EU:n tekoälysäädöksen (AIA)[1] pääkohdat ennakoivasti järjestelmäsi ominaisuuksissa.

4 **Sovi** ulkopuolisten konsulttien kanssa, miten ja missä rajoissa tekoälyä voidaan hyödyntää kehittämisessä. Sovi kirjallisesti laadunhallinnasta ja vastuista.

5 **Ota turvallisuus huomioon** jo tekoälyjärjestelmän suunnitteluvaiheessa, ei vasta käyttöönoton yhteydessä.
- Käytä luotettavia ja testattuja järjestelmiä. Tekoälyjärjestelmät voivat olla monimutkaisia, joten on tärkeää käyttää hyvin suunniteltuja ja testattuja järjestelmiä.

[1] AI Act | Shaping Europe’s digital future (europa.eu)

- Käy läpi järjestelmän toiminta ja asetukset turvallisuuden ja tietosuoja osalta. Rajaa kehitys- ja testausympäristöt vaatimusten mukaisesti.
- Tee ilmoitus kaikista palveluiden käyttöön liittyvistä ongelmista, tunnistamistasi uhkista ja havaitsemistasi riskeistä organisaatiosi ohjeiden mukaisesti.
- Varaudu mahdollisiin ongelmiin. Vaikka tekoälyjärjestelmät ovat usein luotettavia, on tärkeää valmistautua mahdollisiin ongelmiin tai vikoihin.
- Huomioi myös palvelun päättyvä ylläpito ja äkillinen loppuminen. Tämä voi tarkoittaa esimerkiksi varasuunnitelmien tekemistä vaihtoehtoisille tavoille päästä käsiksi ja käsitellä dataa sekä nopeaa viestintää.
- Huolehdi tekoälypalveluita käyttäessäsi kaikista normaaleista palveluiden käyttöön liittyvistä tietoturvakontrolleista (mm. palveluiden asetukset, salasanojen laatu, MFA).

6 **Arvioi ja varmista** tekoälyjärjestelmien turvallisuus ennen käyttöönottoa muun muassa huolellisella testauksella.

7 **Ymmärrä tekoälyn rajoitukset**. Jokainen tekoälyjärjestelmä on suunniteltu tiettyyn tarkoitukseen ja sillä on omat rajoitteensa. Hahmota käyttämäsi tekoälyn tuottamien vastausten vinoumat. Ymmärrä nämä rajoitukset ja sovella sitä ne huomioiden.

8 **Suunnittele** tekoälyjärjestelmän käyttäjäkokemus ihmiskeskeisesti. Siten, että se on helppokäyttöinen ja palvelee käyttäjiensä tarpeita. Tekoälyjärjestelmän tulisi aina palvella ihmisen tarpeita, ei päinvastoin.

9 **Varmista**, että palvelun tuottamat ratkaisut ovat eettisesti hyväksyttäviä, syrjimättömiä ja noudattavat tieto- ja yksityisyydensuojaa. Pyri hyödyntämään netistä löytyviä resursseja, tarkistuslistoja, standardeja, ulkopuolisia tarkastuksia sekä käyttäjien ja sidosryhmien kommentointimahdollisuuksia.

10 **Ota huomioon** tekoälyn vaikutus työntekijöihin. Tekoäly voi muuttaa työpaikan dynamiikkaa ja työntekijöiden rooleja. Pyri tunnistamaan nämä vaikutukset ennakolta ja suunnittele muutokset huolellisesti.

11 **Arvioi**, millainen riski syntyy, jos kaikkea tekoälypalveluun syötettyä organisaation opetusdataa aletaan yhdistelemään, syntyykö tällaisen kasautumisvaikutuksen kautta salassa pidettävää tai muuten sellaista tietoa, joka voi vahingoittaa organisaatiota?

12 **Kouluta** henkilöstöä, jotta kaikilla työntekijöillä on turvalliseen käyttöön tarvittava tieto ja ymmärrys kehitetyn tekoälyjärjestelmän kyvyistä ja rajoista.

13 **Kerro** käyttäjälle selkeästi, mikäli järjestelmää toteutettaessa on hyödynnetty tai se käyttää materiaalien tuottamiseen tai tietojen käsittelyssä tekoälyä (mihin, miten, miksi, mitä palvelua, missä laajuudessa jne.).

14 **Ylläpidä** avointa viestintää. Olipa kyseessä sisäinen tiimi tai ulkoiset sidosryhmät, avoin ja jatkuva viestintä on avain tekoälyn turvalliseen ja tehokkaaseen hyödyntämiseen.

15 **Varmista** organisaation itse tuottamaa tai kehittämää dataa opetusdatana hyödyntäessäsi, että se on laadukasta eikä sisällä esimerkiksi virheellistä, vanhentunutta tai muuta vinoumia aiheuttavia tuloksia. Myös tässä korostuu testauksen merkitys.

ÄLÄ TEE NÄIN:

16 **Älä käytä** tekoälyä ilman tarkoitusta. Se on vain työkalu, jonka käytön pitäisi aina palvella selkeää tarkoitusta. Älä käytä tekoälyä vain sen takia, että se on uusi ja kiinnostava teknologia.

17 **Älä oleta**, että tekoäly ratkaisee (kaikki) ongelmasi. Tekoäly voi auttaa monissa asioissa, mutta se ei ole ratkaisu kaikkeen. Älä oleta, että tekoäly voi korvata kaikki muut työkalut ja prosessit.

18 **Älä unohda** ihmistä tekoälyn takana. Tekoälyjärjestelmät ovat ihmisten suunnittelemia ja toteuttamia, joten huomioi tästä aiheutuvat vinoumat. Huomioi myös ihminen tekoälyjärjestelmän edessä, joka tulkitsee järjestelmän toimintaa ja tuotoksia – järjestelmän taustalla olevien näkökulmien ja tavoitteiden esiintuonti auttaa tässä tulkinnassa.

19 **Älä syötä** palveluun mitään salassa pidettäviä, henkilö- tai turvallisuusluokiteltuja tietoja, mikäli et tiedä onko järjestelmä hyväksytty niitä käsittelemään. Tekoälyjärjestelmissä on huomioitava tietojen ”muistaminen” sekä mahdollinen käyttö järjestelmän kehittämiseen, jolloin tiedot voivat tulla esiin muussa yhteydessä.

- Älä syötä julkiseen palveluun sellaista tietoa, johon sinulla tai organisaatiolla ei ole tekijänoikeutta.
- Älä unohda tietosuojaa. Tekoälyjärjestelmät käyttävät usein suuria määriä dataa, jota on erittäin tärkeää suojata ja käsitellä oikein.

20 **Älä unohda** jatkuvaa seurantaa. Tekoälyn turvallisuus ei ole kertaluonteinen tapahtuma, vaan jatkuva prosessi, joka vaatii säännöllistä seurantaa ja päivitystä.

21 **Älä sivuuta** käyttäjien palautetta. Käyttäjien palaute on arvokas resurssi tekoälyn turvallisuuden ja tehokkuuden parantamisessa. Älä jätä huomiotta heidän kokemuksiaan ja ehdotuksiaan.

3. Tarkistuslista tekoälyn hyödyntämiseen liittyviin kysymyksiin

Olemme koostaneet alle tarkistuslistan asioista, joiden avulla organisaatio voi miettiä tekoälyn hyödyntämiseen liittyviä kysymyksiä ja tarpeita. Osaa niistä on käsitelty osittain aiemmin tässä materiaalissa.

1 Onko tekoälyn käyttö linjassa organisaation strategian ja tavoitteiden kanssa?
- Tekoälyn tulisi tukea yrityksen yleistä strategiaa ja tavoitteita.

2 Miten organisaatio huomioi ihmiskeskeisyyden palveluiden suunnittelussa?
- Palveluun liittyvät sidosryhmät ja käyttäjät tunnistetaan sekä heidät huomioidaan ja heitä kuullaan osana palvelun kehittämistä ja sen elinkaaren hallintaa.

3 Onko olemassa selkeä tarve tai ongelma, jonka tekoäly voi ratkaista?
- Tekoälyn ei tulisi olla ratkaisu ongelman etsimiseen, vaan sen tulisi ratkaista olemassa oleva tarve tai ongelma. Yleisiä kohteita tekoälyn käyttöön on erilaisten työvaiheiden automatisointi, jolloin pyritään parantamaan nopeutta, tehokkuutta, laatua tai kaikkia näistä.

4 Onko organisaatiolla selkeä käsitys siitä, miten tekoäly toimii ja miten sitä tulisi hallita?
- On tärkeää ymmärtää tekoälyn periaatteet ja etenkin sen toiminnan rajoitukset. Näihin liittyvät myös kontekstit, tietojärjestelmät ja alustat, prosessit ja palvelut, jossa tekoälyä käytetään, sille ne vaikuttavat odotuksiin, tulkintaan ja vaatimuksiin.

5 Onko tekoälyyn liittyviä eettisiä- ja vastuullisuuskysymyksiä otettu huomioon?
- Nämä kysymykset voivat liittyä esimerkiksi syrjintään, läpinäkyvyyteen ja tietosuojaan. Välittömien kysymysten lisäksi tulee pohtia pitkän aikavälin sekä suuren skaalautuvuuden (määrien) merkityksiä, jolloin pienetkin asiat voivat kasvaa merkittäviksi.

6 Miten palveluiden tietoturvallisuudesta on huolehdittu?
- Tekoälypalveluiden turvallisuutta tulisi hallita samalla tavalla kuten kaikkien organisaation tuottamien tai käyttämien palveluiden täydennyttynä tekoälypalveluiden ainutkertaisuuteen ja erilaisuuteen liittyvillä vaatimuksilla. Varmista, että tekoälypalvelun tarjoaja on huolehtinut tietoturvallisuudesta palvelussaan. Vaikka tiedot olisivat julkisia, joku saattaa haluta esimerkiksi vaikuttaa tuloksiin tai väärentää tietoja. Arvioinnin tulisi olla osa organisaatioiden palveluiden kehittämiseen tai käyttöönottoon liittyvää riskienarviointia, jossa huomioidaan myös henkilötietojen käsittelyssä tarvittavat riski- ja vaikutustenarvioinnit.

7 Tunnista kokonaisuuteen liittyvä sääntely
- Noudata tekoälypalveluiden käytössä kaikkia soveltuvia säädöksiä ja lakeja. Niihin voi sisältyä EU-tasoista tai kansallista sääntelyä, tietosuojalakeja, teollisuuden standardeja ja muita sääntöjä. Löydät esimerkkejä näistä myöhemmin tässä materiaalissa. Kannattaa muistaa, että esimerkiksi EU-alueella merkittävässä roolissa oleva AI

Act on vasta juuri hyväksytty toukokuussa 2024 [2], ja sen toimeenpano tulee kestämään kokonaisuudessaan vuosia.[3]

8 Selvitä vastuukysymykset ennakkoon
- Tekoäly ei ole vastuussa virheistä - ihmiset ovat. Ole aina valmis ottamaan vastuu tekoälyjärjestelmän päätöksistä ja toimista.

9 Varmista tuki myös tekijänoikeus- ja muissa laillisuuskysymyksissä
- Selvitä etukäteen, millaisia tekijänoikeuksiin liittyviä haasteita tuottamiesi tietojen jakamiseen liittyy ja varmista tarvittavan lainopillisen tuen saatavuus mahdollisissa ongelmatilanteissa.

10 Onko tekoälyn hyödyntämisessä huomioitu esimerkiksi tiedonhallintaan, julkisuuslakiin ja hyvään hallintoon liittyvät velvoitteet?
- Palveluiden käyttöönotossa tulee suorittaa arviointi palvelun vaatimustenmukaisuuden toteutumisesta, jonka yksi osa on lainsäädäntövelvoitteiden täyttyminen. Ota huomioon myös tekoälypalveluiden, jotka ovat usein pilvipalveluita, tuottamisessa huomioitavat velvoitteet. Tunnista myös mahdollisuus tarvittaessa käyttää rajoitettuja, itse kansallisesti tuotettuja tekoälypalveluita, joiden toiminta pohjautuisi esimerkiksi vain itse palveluun vietyyn opetusdataan.

11 Onko tekoälyn käyttöön liittyviä riskejä arvioitu ja onko riskienhallintasuunnitelmaa?
- Tekoälyn käyttöön liittyy useita riskejä, mukaan lukien teknologiset riskit, liiketoimintariskit ja maineeseen liittyvät riskit. Onko riskit tunnistettu ja millaisilla keinoilla niitä hallitaan? Entä miten näiden riskien toteutuminen on otettu huomioon jatkuvuus-, varautumis- ja kriisiviestintäsuunnitelmissa?

12 Onko organisaatiolla resursseja ja osaamista tekoälyn käyttöönottoon ja ylläpitoon sekä koulutukseen?
- Tekoälyn käyttöönotto ja ylläpito vaativat sekä teknistä että substanssiin liittyvää osaamista ja käyttäjien koulutusta. Tarvittava käyttäjäkoulutus tekoälypalveluiden käyttäjäksi tulee olemaan jatkossa osa ICT:n käyttötaitoja, vastaavasti palveluiden kehittäjien osalta tarvitaan kokonaan uudenlaista, tekoälypalveluiden erityispiirteisiin liittyvää osaamisen kehittämistä.

13 Miten tekoälyn käyttö vaikuttaa työntekijöihin ja asiakkaisiin?
- Tekoälyn käyttöön voi liittyä sosiaalisia ja kulttuurisia vaikutuksia, kuten työtehtävien muutoksia ja asiakkaiden hyvinkin vaihtelevaa suhtautumista tekoälyyn.

14 Onko organisaatiolla suunnitelmaa tekoälyn käytön arvioimiseksi ja seuraamiseksi?
- On tärkeää arvioida tekoälyn vaikutuksia säännöllisesti ja tehdä tarvittavia muutoksia sen käyttöön. Vaikuttavuuden ja muiden mitattavien asioiden osalta tulee luoda mittarit etukäteen, joita säännöllisesti mitataan ja tulosten perusteella toimintaa kehitetään.

[2] https://www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/ - 21.5.2024
[3] https://data.consilium.europa.eu/doc/document/PE-24-2024-INIT/en/pdf - 14.5.2024

15 Miten tekoälyn käyttäminen sovelluskehityksessä sallitaan?
- Millainen politiikka organisaatiolla on tekoälyn hyödyntämiseen sovelluskehityksessä? Tämä on arvioitava riippumatta siitä, kehittääkö organisaatio itse sovelluksia vai ei. Mikäli tekoälypalveluiden käyttö sallitaan sovelluskehityksessä, mitkä ovat ne ehdot, rajoitukset ja kontrollit, joita siinä edellytetään? Miten varmistetaan, että salassa pidettävää tietoa (myös mahdolliset algoritmit) ei vuoda ulkopuolisille toimijoille?

16 Miten organisaatio on päivittänyt tekoälyn hyödyntämiseen tai käyttämiseen liittyvät vaatimukset osaksi hankinta-ohjeita ja kilpailutukseen liittyviä vaatimuksia?
- Tekoälypalveluita tuottavien tai hyödyntävien palveluiden julkinen hankinta on vielä varsin uutta ja siinä tulee ottaa huomioon useita tässä materiaalissa jo esille nostettuja tehtäviä. Koska tekoälypalvelut toimivat ”black box” eli ”musta laatikko” -periaatteella, eri toimittajien palveluiden vertaileminen ilman niiden oikeaa käyttökokemusta pelkästään kirjallisesta dokumentaatiosta voi olla haasteellista. Kilpailutukseen ja hankintavaatimusten laatimiseen vaikuttaa merkittävästi se, millaiseen käyttötarpeeseen palvelua ollaan hankkimassa.

3.1 Muuta huomioitavaa

1 **Tekoäly luo valtavia mahdollisuuksia, mutta se tuo mukanaan myös merkittäviä riskejä**. Verkkorikolliset, valtiollisen tason toimijat ja muut tahot pystyvät hyödyntämään sitä meitä helpommin välittämättä lainsäädännöstä tai toiminnan eettisyydestä. Toisaalta samoja menetelmiä, joita käytetään näiden hyökkäysten laatimiseen, voidaan ainakin osin käyttää niitä vastaan puolustautumiseen.

2 Tekoälyn hyödyntäminen tulee **merkittävästi laajentumaan** siitä, miten kuvittelemme sitä käytettävän tänä päivänä. Esimerkiksi Internet-verkkoa käytettiin 1990-huomattavasti rajatummin, mitä 2000-luvulla älylaitteiden yleistymisen myötä. Sama on odotettavissa tekoälyn hyödyntämisen osalta. Jos jonkin teknologisen palvelun tai ilmiön yleistymiseen on kulunut 1990-2000-luvun alusta esimerkiksi kymmenen vuotta, vastaavanlainen loikka saadaan tehtyä nyt huomattavasti nopeammassa ajassa, jopa vuosissa tai käyttötapauksesta riippuen, kuukausissa. Tämä johtuu teknologisen suorituskyvyn merkittävästä, ns. eksponentiaalisesta kehittymisestä vuosittain.

3 Tekoälyn käytöstä **on useita hyötyjä**, mutta sen vastapainona se myös **kuluttaa resursseja ja lisää riskejä**. Ennen käyttöönottoa tulee tehdä riittävät hyödyllisyys-, vastuullisuus- ja riskiarviot. **Vastuullisessa** tekoälyn hyödyntämisessä huomioidaan myös, onko tekoälyn käyttö kaikissa käyttötapauksissa oikeasuhtainen valinta. IEA:n raportin mukaan datakeskusten, tekoälyn ja kryptovaluuttasektorin sähkönkulutus voi kaksinkertaistua vuoteen 2026 mennessä [4].

4 Riskitasoa arvioitaessa **turvallisin tapa** kokeilla tekoälyä on silloin, kun **onnistuminen luo paljon positiivista** tai on todella hyödyksi, mutta epäonnistuminen ei juurikaan tuota vahinkoa. Mikäli tekoälyn tekninen ja hallinnollinen toteutus (esimerkiksi neuroverkko) muodostaa vaikeasti hallittavan ja hahmotettavan ”mustan laatikon”,

[4] https://www.iea.org/reports/electricity-2024/executive-summary

riskit on hallittava prosessissa sitä ennen (oikea kohdistus, data ja sen laatu) ja sen jälkeen (tuotosten valvonta, filtteröinti ja tehokas virheiden korjaus sekä oppiminen).

5 Tekoälyllä **ei ole suunnitelmaa tai pyrkimystä oikeaan tulokseen**, väärin alkanut polku päätelmiä voi johtaa yhä vain huonompaan tulokseen.

6 **Tekoäly antaa tilastollisesti oikean vastauksen**, ei oikeaa vastausta; sen tuottama vastaus on niin hyvä tai kun huono kuin on sen opetuksessa käytetty data. Tämä koskee myös organisaation itse tuottamaa opetusdataa, sen käyttäminen ei automaattisesti takaa oikeita tuloksia, vaan on riippuvainen esimerkiksi tiedon oikeellisuudesta ja ajantasaisuudesta.

7 Tekoäly on **tehokas tuottamaan vastauksia**, jolloin **virheet monistuvat nopeasti** ja tämä tulee huomioida kehitettäessä tekoälyä hyödyntävää palvelua tai prosessia. Näissä on samaan aikaan kehitettävä vastaavasti riittävän tehokkaita ja skaalautuvia virheiden korjaus- ja hallintakeinoja. Näihin lukeutuu myös selkeä toimintatapa tekoälyn käytön pysäyttämiseksi, mutta tämä ei vielä ratkaise mahdollisten aiheutuneiden vahinkojen minimointia ja korjausta.

8 Tulevaisuudessa **opetusdatan manipulointi** tulee olemaan kasvava riski, joten tekoälyn tuottaman tiedon laatua tulee valvoa.

4. Mistä voit opiskella lisää?

Voit luonnollisesti hyödyntää tekoälypalveluita oman osaamisen kehittämiseksi pyytämällä sitä selittämään asioita, mutta alla muutama tunnettu tukimateriaali ja koulutus aiheeseen liittyen:

- Kansalliseen AuroraAI tekoälyohjelmaan liittyvä webinaarisarja Haus Kehittämiskeskuksen eOppiva-palvelussa:
  - Osa 1. Mistä tekoälyssä on kyse?
  - Osa 2. Tekoälyn toimintaperiaate
  - Osa 3. Tekoäly yhteiskunnassa

- Kehittäjän opas – Tekoälyn vastuullinen hyödyntäminen (DVV)
- Helsingin yliopiston laatima maksuton Tekoälyn etiikka -kurssi (mooc.fi)
- Algoritminen syrjintä ja yhdenvertaisuuden edistäminen: Arviointikehikko syrjimättömälle tekoälylle (VNK)
- Tekoälyn eettinen ohjeistus, huoneentaulu (VM)
- Turvallisen tekoälykehittämisen opas (DVV)
- Ihmisten tekoäly | Tiedekulma | Helsingin yliopisto (helsinki.fi)
- Elements of AI-verkkkurssi: https://www.elementsofai.com/fi/
- EU:n digisäädökset (VM)
- Laki digitaalisten palvelujen tarjoamisesta
- Automaattista päätöksentekoa koskevan hallinnon yleislainsäädännön valmistelu
- Koneoppiminen digitaalisen turvallisuuden teknisessä valvonnassa (VM)
- FCAI
- Etairos
- OECD:n tekoälyportaali
- WEF: Artificial Intelligence for Children
- Euroopan parlamentin tutkimuksia tekoälystä
  - esimerkiksi Artificial intelligence and cybersecurity - Artificial intelligence and cybersecurity | Epthinktank | European Parliament
- Forecasting potential misuses of language models for disinformation campaigns and how to reduce risk (2023) openai.com/research/forecasting-misuse
- Disrupting deceptive uses of AI by covert influence operations | OpenAI (30.5.2024)
- Tekoälyn soveltamisen kyberturvallisuus ja riskienhallinta (2021) traficom.fi/ajankohtaista/julkaisut
- Tunnisteet ja tietosuoja – Anonymisointi ja sen rajat (2021) kyberturvallisuuskeskus.fi/fi/ajankohtaista/ohjeet-ja-oppaat/ohjeet-ja-oppaat-tietoturva-ammattilaisille
- Tekoälyn hyödyntäminen vihapuheen seurannassa (2021) [tapausesimerkki] julkaisut.valtioneuvosto.fi/
- Tekoäly viranomaistoiminnassa – eettiset kysymykset ja yhteiskunnallinen hyväksyttävyys (2019) julkaisut.valtioneuvosto.fi/

- Ohjelmistorobotiikka ja tekoäly – soveltamisen askelmerkkejä (2018) julkaisut.valtioneuvosto.fi/ (osa vanhentunutta, mutta vastaa edelleen useiden toimijoiden maturiteettia
- Standardeja ja viitemateriaalia
  - ISO/IEC 23894:2023, Guidance on risk management
  - ISO/IEC 22989:2022, Artificial intelligence concepts and terminology
  - ISO/IEC 38507:2022, Governance implications of the use of artificial intelligence by organizations
  - ISO/IEC TR 24027:2021, Bias in AI systems and AI aided decision making
  - ISO/IEC TR 24028:2020, Overview of trustworthiness in artificial intelligence

- Muuta materiaalia
  - DVV:n digiturvan kesäseminaari 15.6: Tekoälyn mahdollisuudet ja julkisen sektorin digiturvan nykytila, jossa tämän materiaalin versio 1.0 julkaistiin, sisältää useamman tekoälyä käsittelevän esityksen – linkki tallenteeseen
  - VAHTI:n AI-webinaari (12.9.2023)
  - DVV:n digiturvan seminaari 12.12.2023.
    - Tekoäly arjessa – miten käytät tekoälypalveluita helposti ja turvallisesti?
    - Päivitetyt VAHTI-verkoston hyvät käytännöt tekoälyn hyödyntämiseen
    - Viisi myyttiä tekoälystä ja kuinka ne murretaan
    - Miten taklata tekoälyn hyödyntämisen haasteet?
      Linkki tallenteeseen.
  - DVV:n, KELA:n ja VERO:n AI-tapahtuma (1.9.2023)
  - Tekoälyn vinoumien välttämisen hankkeen paneelikeskustelu (23.8.2022) tuotettu tarkistuslista
  - ”Tekoälyä ollaan jo sääntelemässä”, VM:n blogi (28.6.2023)
  - AIGA-hankkeessa kehitetty ”tiimalasimalli” esimerkkinä tekoälyn eri tasojen hallintakokonaisuudesta organisaatioissa sekä elinkaarimalli, joka listaa tehtäviä (2022)
  - AI:ta pohjustavasta hyvästä datanhallinnasta webinaaritallenne; Loihde (8.11.2023)
  - Kun jokainen päivä voi olla aprillipäivä - Mistä deepfakeissa on kysymys? 1.4.2024 linkki tallenteeseen.