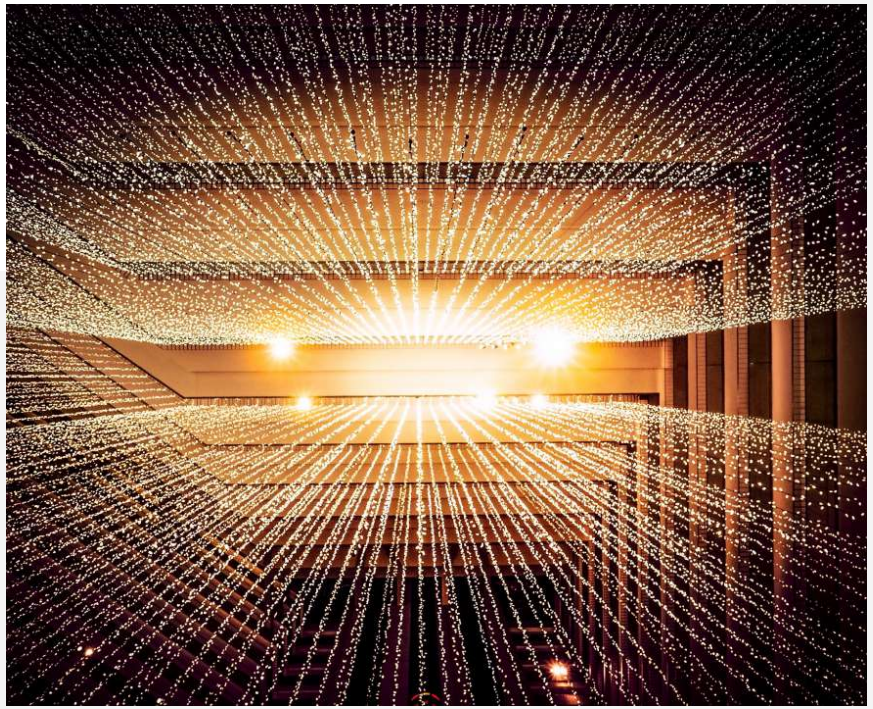
ETAIROS



KÄYNNISTYSKOKOUS

7.10.2019 KLO 13-16

Tapahtumatalo Bank, Helsinki

@etairosai

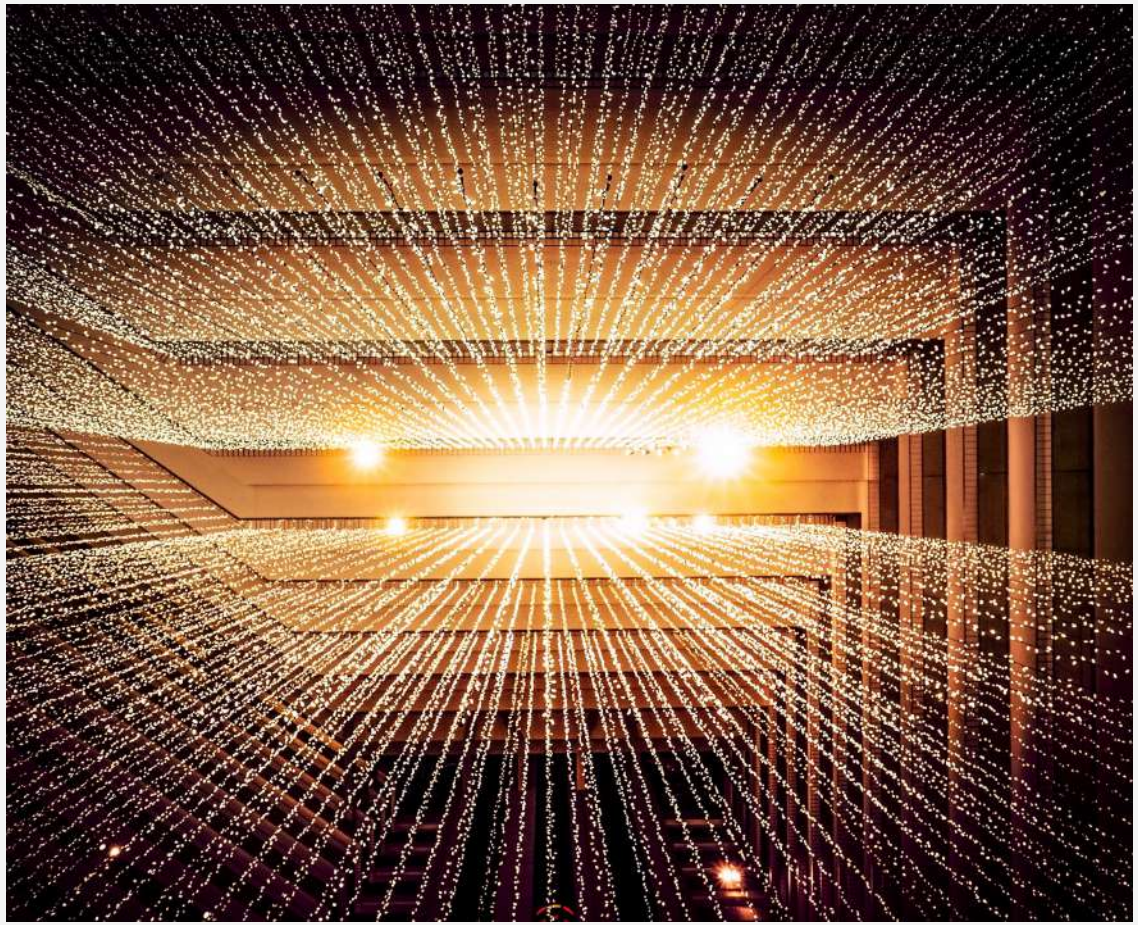
SUOMEN AKATEMIA

strateginen TUTKIMUS

Ohjelma 7.10.2019

13.00 - 13.10	Tervetuloa *Tutkimushankkeen vastuullinen johtaja, Jyrki Nummenmaa, Tampereen yliopisto*
13.10 - 13.35	ETAIROS lyhyesti *Jyrki Nummenmaa, Tampereen yliopisto*
13.40 - 14.00	Tekoälyn yhteiskunnalliset ja eettiset ulottuvuudet *Mika Nieminen, VTT* *Jaana Hallamaa, Helsingin yliopisto*
14.00 - 14.20	Kahvitauko
14.20 - 14.50	ETAIROS tapaustutkimukset ja teema-alueet *Jaana Leikas, VTT; Mika Pyyhtiä, Megical; Sauli Jäppinen, Kela*
14.50 - 15.30	Ryhmätyöskentely: Tarkennuksia, odotuksia ja tavoitteita ETAIROS-työlle
15.30 - 15.55	Ryhmätöiden purku *Kalle Piirainen, 4FRONT*
15.55 - 16.00	Päivän yhteenveto sekä seuraavat askeleet *Jyrki Nummenmaa, Tampereen YO*

ETAIROS lyhyesti



Hankkeen tausta, tarkoitus ja kysymykset

- Tekoäly-teknologioiden (AI) ja autonomisia piirteitä sisältävien järjestelmien käyttö lisääntyy ja kehittyy voimakkaasti
- Tekoälyn käytön ongelmat liittyvät mm tasapuolisuuteen, turvallisuuteen, hyväksyttävyyteen, yksityisyyteen, jne.
- Ongelmien syitä ovat mm.
  - järjestelmien sääntökokonaisuuksien monimutkaisuus,
  - koneoppimisen opetusdatan vinoutumat ja
  - yleisemmin se, että eettisiä ongelmia ei ymmärretä eikä huomioida.

Keskeiset ohjaavat kysymykset

- Kuinka tunnistetaan tekoälyn käytön eettiset riskit?
- Kuinka suunnitella eettisesti hyväksyttäviä tekoälyjärjestelmiä?
- Kuinka käyttää tekoälyjärjestelmiä eettisesti hyväksyttävästi?
- Kuinka tekoäly muuttaa yhteiskuntaa ja etiikkaa?
- Kuinka säädellä tekoälyjärjestelmien käyttöä?

Tekoälyteknologiat

Tekoäly (AI)

Teknologiat

Kielen-/puheentunnistus,
Konenäkö,
Hahmontunnistus
…

Koneoppiminen, syväoppiminen,
neuroverkot,
…
Sääntöpohjainen tekoäly (AI)

Tekniset implementaatiot

Yksittäiset tekniset
ratkaisut ja alustat

Sovelluksia

Big Data,
analyysi,
Optimointi
…

Tutkimuksen konkreettiset tavoitteet

- Tunnistetaan tekoälyjärjestelmien eettiset ongelmat (jo suunnitteluvaiheessa)
- Kehitetään suunnittelumenetelmiä, joilla nämä ongelmat vältetään, mm.
  - Roolitetaan tekoäly (AI) ja ihminen järjestelmissä oikein
  - Analysoidaan ja simuloidaan malleja järjestelmistä
  - Muokataan koneoppimismenetelmiä
  - Identifioidaan korrektit sovelluskohteet ja -tavat
  - Etsitään tarpeita ja ratkaisuita yhteiskunnan säätelyyn tekoälyä hyödynnettäessä

4front.fi

Hankekonsortion esittely

Hankkeen toteutukseen osallistuvat

- Tampereen yliopisto
- Helsingin yliopisto
- VTT Oy
- Turun yliopisto
- Jyväskylän yliopisto
- 4FRONT Oy

Vastuuhenkilöt

- Hankkeen vastuullinen johtaja Jyrki Nummenmaa, professori, Tampereen yliopisto
- TP1: Tekoälyn tulevaisuus ja ennakointi
  Nadezhda Gotcheva, erikoistutkija, VTT
- TP2: Toiminnan etiikka ja riskit
  Jaana Hallamaa, professori, Helsingin yliopisto
- TP3: Eettinen suunnittelu
  Pertti Saariluoma, professori, Jyväskylän yliopisto
- TP4: Koneoppiminen ja etiikka
  Jaakko Peltonen, professori, Tampereen yliopisto
- TP5: Tekoälyn eettisen hyödyntämisen koordinaatio ja ohjaus, Jaana Leikas, dosentti, johtava tutkija, VTT
- TP6: Vuorovaikutus, Kalle A. Piirainen, dosentti, LUT ja vanhempi asiantuntija, osakas, 4FRONT

Rahoittaja

SUOMEN AKATEMIA

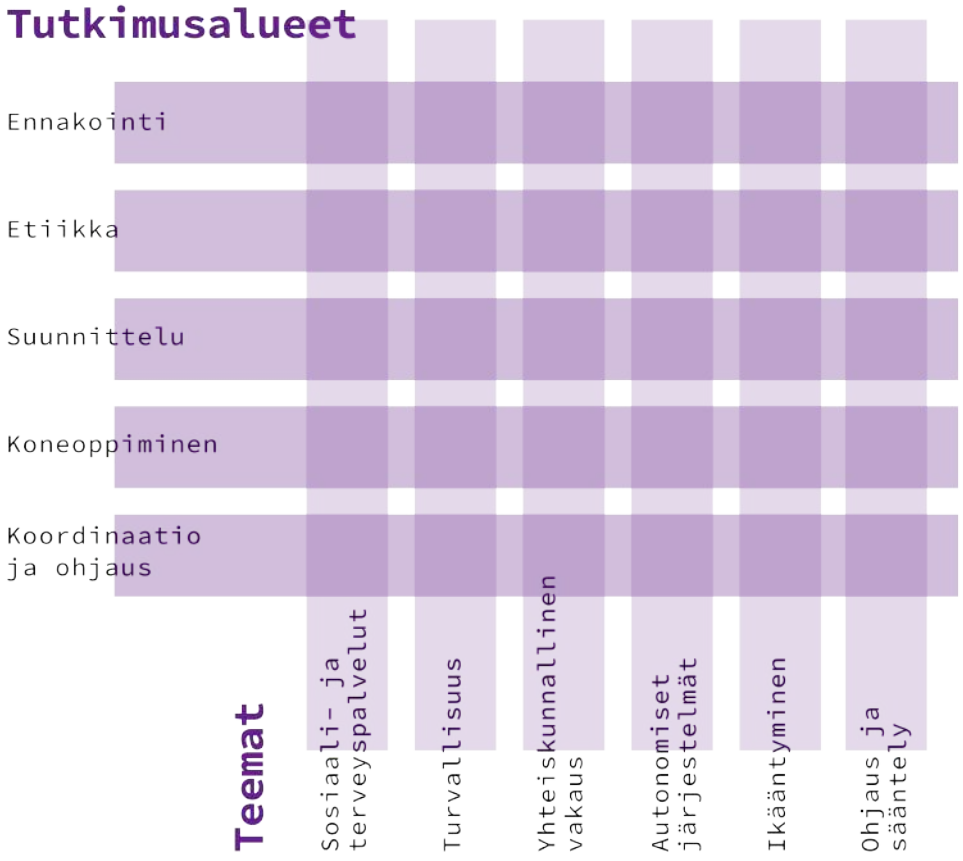
strateginen TUTKIMUS

ETAIROS

Hankkeen rakenne: tutkimusalueet ja teemat

Hankkeen looginen työvuo rakentuu

- Tutkimusalueiden (työpakettien) ketjuuntumiselle ja
- Näiden tulosten soveltamiselle, kokeilulle ja arvioinnille,
- Yhdessä kumppaneiden ja työpakettien välillä tapahtuvassa vuorovaikutuksessa, ja
- teema-alueilla tapahtuvassa yhteiskehittämisessä tapauskumppaneiden kanssa.



Vuorovaikutusfoorumit

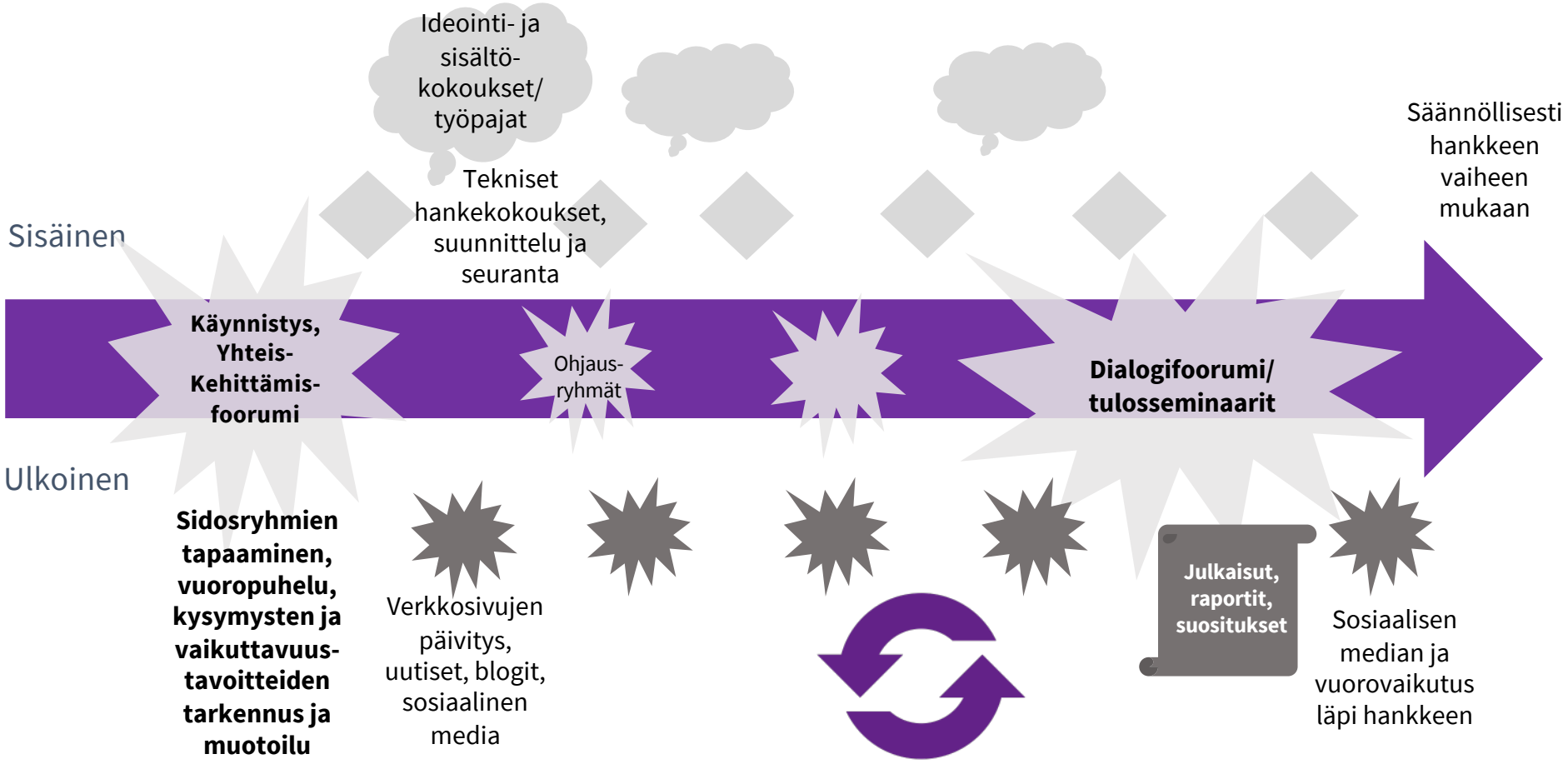
Yhteiskehittämisfoorumit

- Hankkeen kumppaneille
- Kahdesti vuodessa (loka-marraskuu ja huhti-toukokuu)
- Keräävät teema-alueet ja käyttötapaukset yhteen
- Hankkeen sisältöjen kehittäminen ja palaute
- Dialogi tulosten merkityksestä kumppaneille ja ideoiden jatkokehittely

Dialogifoorumit

- Avoin kaikille sidosryhmille
- Kerran vuodessa (alkuvuodesta, hankkeen vaiheen mukaan)
- Teema-alueiden sisällöllisten havaintojen ja tulosten esittely
- Dialogi tulosten merkityksestä laajemmin ja ideoiden jatkokehittely

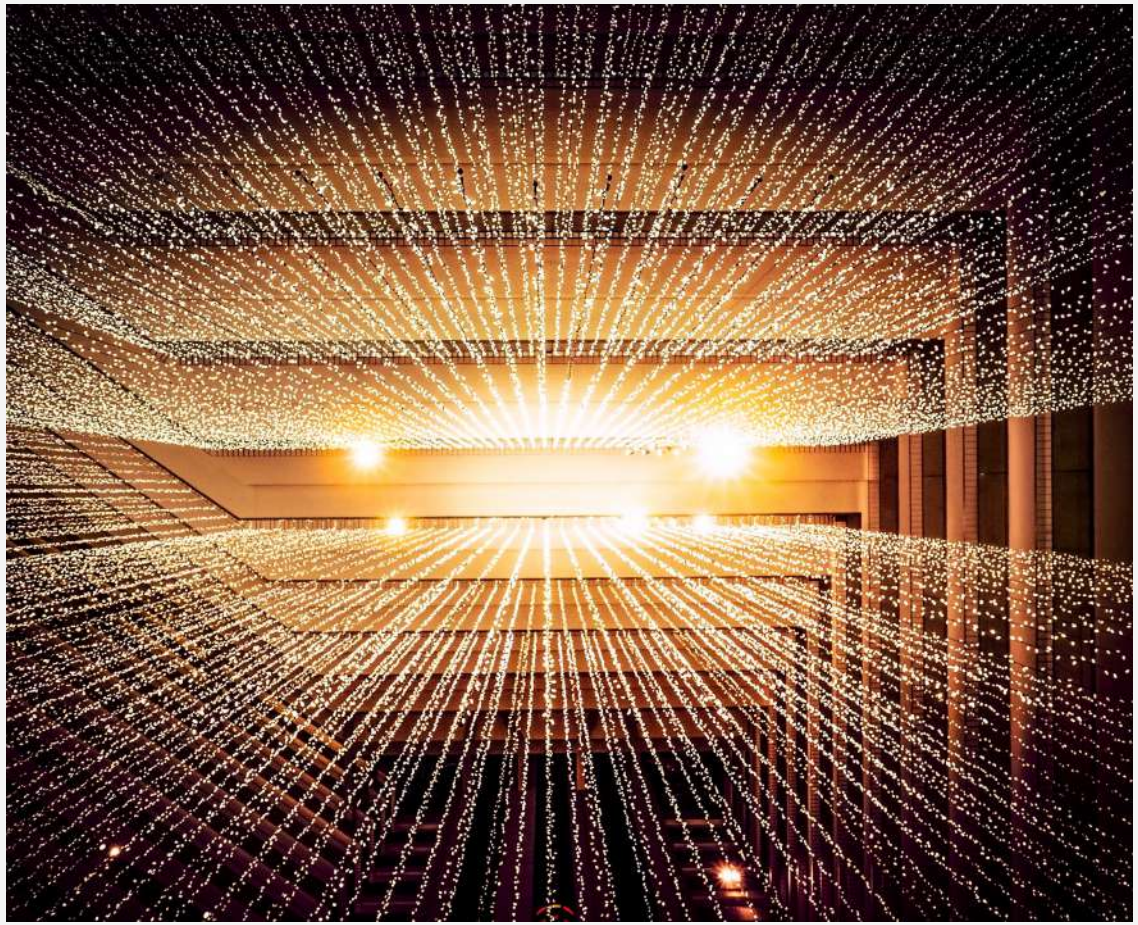
Vuorovaikutuksen ensimmäinen puoli vuotta



Yhteiskehittäminen;
Tapaamiset, esitykset, haastattelut, kyselyt ja muu aineiston keruu

Sisältötyön ja kumppaneiden tarpeen/halun mukaan projektin läpi

Tekoälyn yhteiskunnalliset ja eettiset ulottuvuudet



Tekoäly yhteiskunnassa

- Käynnissä syvällekäyvä sosio-tekninen muutos tai jopa murros yhteiskunnassa: Instituutiot, yhteisöt, organisaatiot, yksilöt…

- Kuten, polttomoottori, sähkövoima…?

Tekoäly yhteiskunnassa

- Teknologia ei deterministisesti määrittele yhteiskunnallisen muutoksen suuntaa. Ihmiset tekevät sen valinnoillaan ja päätöksillään kehittäessään ja hyödyntäessään teknologiaa.
- Melvin Kranzberg (1986): ”Teknologia ei itsessään ole hyvä tai paha; mutta se ei ole myöskään neutraali.”

Tekoäly yhteiskunnassa

- Tekoäly poikkeaa mekaanisesta teknologiasta yhteiskunnalliselta luonteeltaan.
- **AI-teknologia on sosiaalisesti ja kulttuurisesti ”puitteistunut”**: Ohjelmointi ja data ovat ihmisten tuottamia ja kantavat mukanaan inhimillisiä merkitysrakenteita, virheitä jne.
- Sen hyödyntämisellä on **välittömiä, suhteellisen nopeasti nähtävissä olevia tai välillisiä ja pitkän-ajan kuluessa toteutuvia sosiaalisia, poliittisia ja organisatorisia vaikutuksia.**
- Antropomorfismin vaara: Minkälainen ”toimija” tekoäly on?

Tekoäly yhteiskunnassa

- AI-teknologian hyödyntäminen on myös aina johonkin erityiseen sosiaaliseen ja organisatoriseen ympäristöön liittyvää.
- Siksi konkreettisia yhteiskunnallisia vaikutuksia on tarkasteltava erilaisissa sosiaalisissa konteksteissa kokonaisvaltaisesti yksilön, yhteisön, organisaatioiden ja laajemmin yhteiskunnan kannalta.
- Yhteiskunnallisuus merkitsee myös sitä, että teknologian kestävä kehittäminen ja käyttöönotto edellyttää dialogisuutta: kehittäjien, käyttäjien, niiden tahojen ja ihmisten joihin vaikutukset kohdistuvat keskustelua käyttöönotosta ja sen vaikutuksista omista arvolähtökohdistaan.

Tekoäly yhteiskunnassa

- Mitkä ovat yhteiskunnalliset riskit?
  - AI voi johtaa älyttömiin päätöksiin
  - AI voi syrjiä ihmisiä/-ryhmiä
  - AI voi muuttaa perustavanlaatuisesti sitä miten olemme maailmassa

Emmekä edes välttämättä huomaa mitään näistä!

ETAIROS

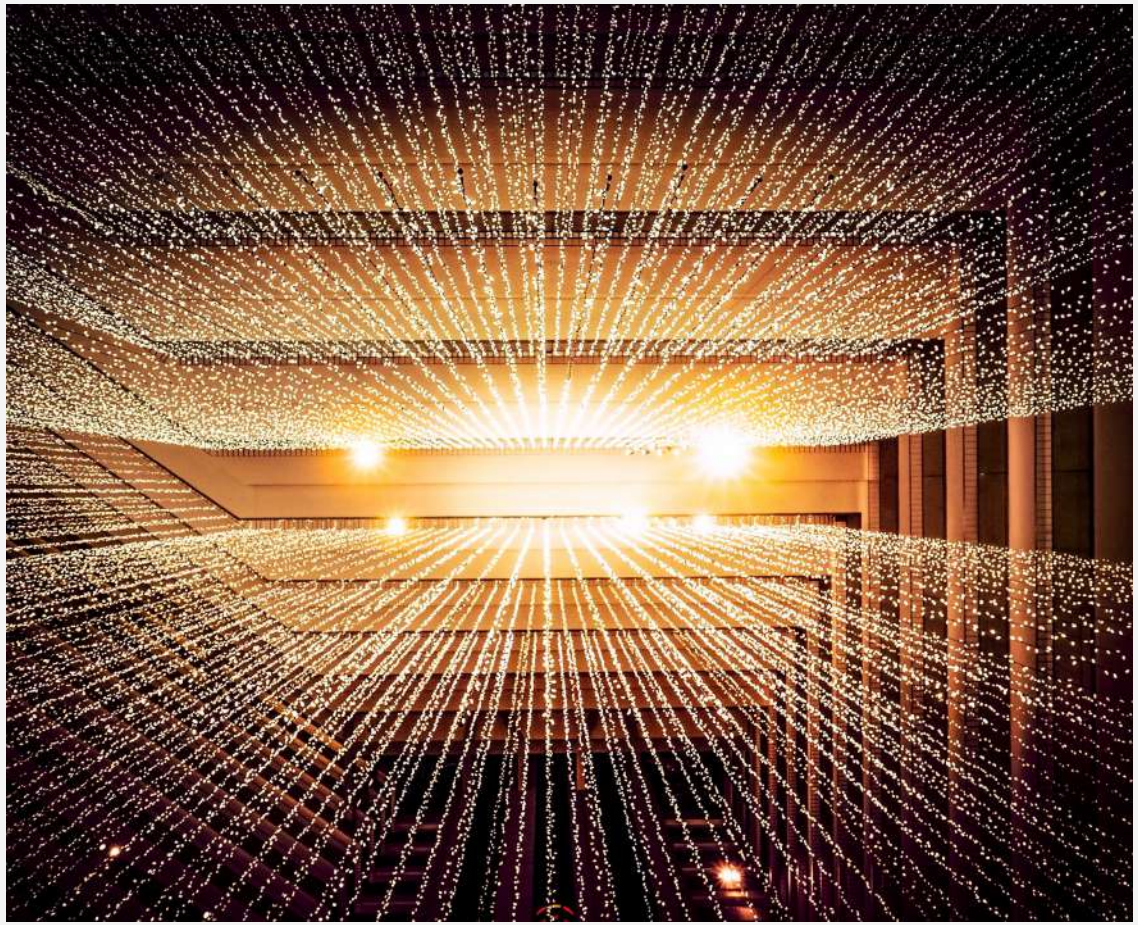
ETAIROS

- ET – AI - ROS
  - Eettistä tekoälyä julkisten hankkeiden palvelukseen
- **ἑταῖρος**
  - Taustalla indoeurooppalaisen kantakielen itseä tarkoittava sana (swe) ja siitä johdettu merkitys: oman heimon jäsen.
  - Kreikassa
    - Kumppani, partneri, ystävä
  - Sanaan vähitellen yhdistyvä negatiivinen sivumerkitys
    - Yhteistyösuhteessa ahne ja kateellinen ystävä (Matt. 20:13)
    - Epäkunnioittava ja piittaamaton ystävä (Matt. 22:12)
    - Petollinen ystävä (Matt. 26:50)

Tekoälyjärjestelmät toiminnan laajennuksena

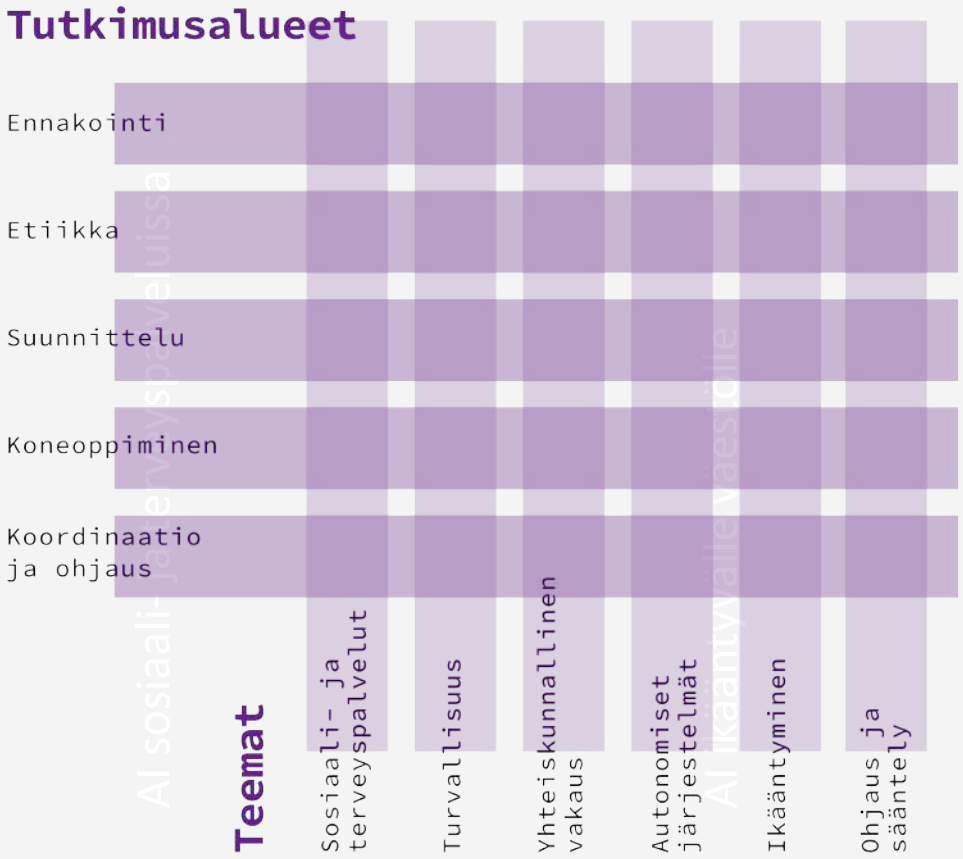
- Etiikan näkökulma:
  - Mitä hyvää toiminnalla tavoitetaan?
  - Ketä toiminta hyödyttää?
  - Millaisia kustannuksia ja haittoja aiheutuu? Kuka maksaa ja korvaa ne?
  - Mitä vahinkoa aiheutuu ja ketkä siitä kärsivät?
  - Millaisiin epäonnistumisiin ja virheisiin toiminta altistaa?
- Tavoitteena selvittää:
  - Kuinka hyvää voidaan tehdä ja edistää sekä pahaa välttää ja ehkäistä tekoälyjärjestelmien ja ihmisten yhteistoiminnalla.

Tapaustutkimukset ja teema-alueet



Teema-alueet, työpaketit ja käyttötapaukset

- Sidosryhmien suuntaan työ jäsentyy teema-alueiden ja tapauksien kautta
  - Työpakettien teoreettinen työ jalkautuu yhteiskehittämiseksi teema-alueilla ja erityisesti käyttötapauksissa.



ETAIROS AI:n temaattiset käyttötapaukset (1/3)

YHTEISKUNNALLINEN VAKAUS

- Tekoäly turvallisuuden hallinnassa ihmisen partnerina
- Reaaliaikaista dataa käyttävä, oppiva, tilannekuvaa jatkuvasti muodostava tekoälyjärjestelmä
- Eettisyyden huomiointi suunnittelussa, eettisesti epäilyttävien tilanteiden tunnistaminen käytössä

AUTONOMISET JÄRJESTELMÄT

- Tekoälyn ja ihmisen vuorovaikutus ja sen vaikutus esim. turvallisuuteen ja päätöksentekoon
- AI liikennejärjestelmien integroinnissa
- Liikenne palveluna (MaaS) –tematiikka

KANSALLINEN TURVALLISUUS

- Säädösten ja toimintamallien riittävyys sekä sosio-tekninen uudistaminen
- Sääntöpohjaisuuden vahvistaminen
- Kompleksisten uhkien tunnistaminen ja hallinta

ETAIROS

ETAIROS AI:n temaattiset käyttötapaukset (2/3)

SOSIAALI- JA TERVEYSPALVELUT

- Yhteistoimijuutta tuottavat tekoälyjärjestelmät; tavoitteena asiakkaiden ennaltaehkäisevä, terveyttä ja hyvinvointia ylläpitävä toiminta
- Yhteistoimijuutta tuottava tekoälymalli, jonka varassa voidaan toteuttaa eettisesti kestäviä, tekoälypohjaisia sote-sovelluksia

IKÄÄNTYMINEN

- AI:n käyttötapaukset ja eettiset kysymykset kotihoidossa
  - Asiakas / omainen / hoitohenkilökunta /julkinen palveluntuottaja
- Eettisyys AI-ekosysteemeissä & AI:n hyödyntämisen ohjauskeinot
- Datan hallinta (datalähtöinen etiikka & tietoinen suostumus)
- AI päätöksenteossa

ETAIROS AI:n temaattiset käyttötapaukset (3/3)

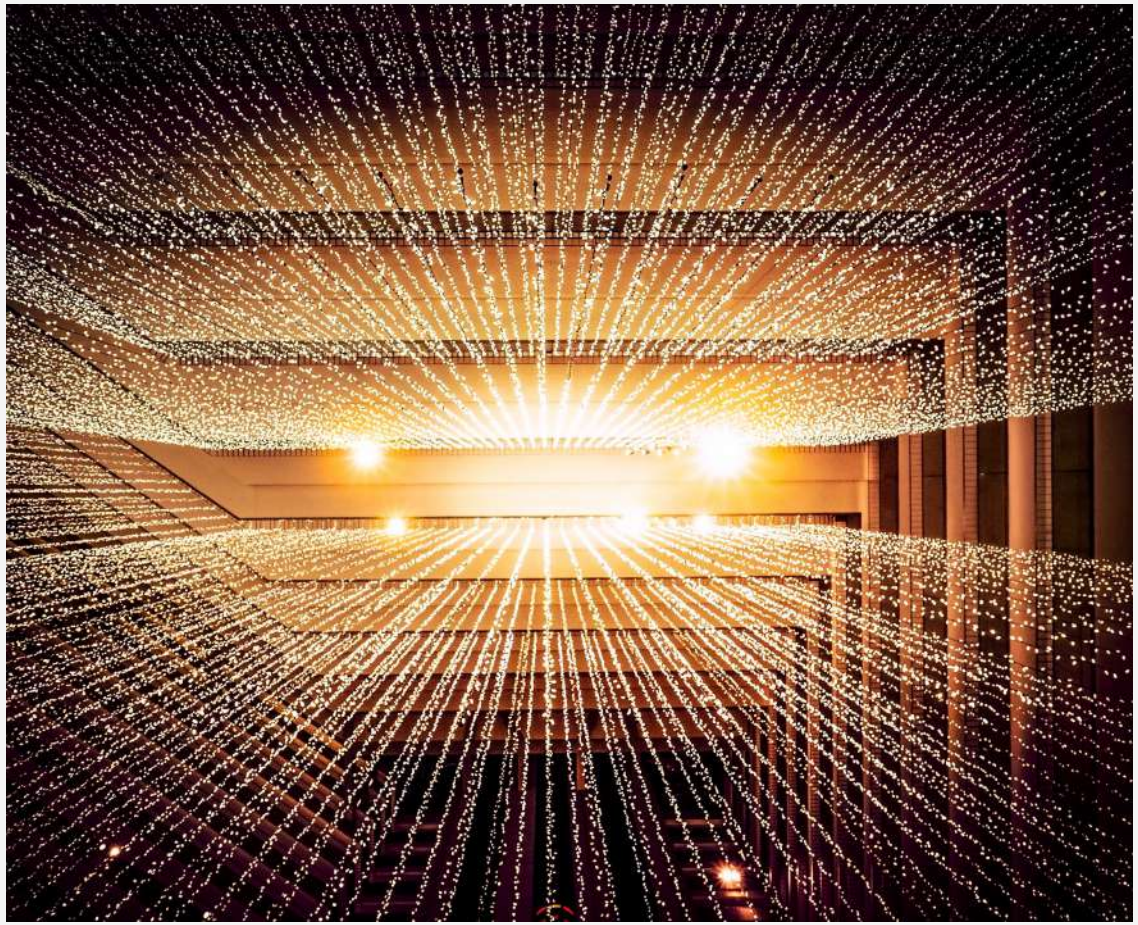
KOKONAISVALTAINEN OHJAUS JA SÄÄNTELY

- 5 ulottuvuutta:
  - Datalähtöisyys (hallinta datan käytössä ja hyödyntämisessä)
  - Oikeudellisuus (oikeudelliset reunaehdot)
  - Julkinen politiikka
  - Etiikkalähtöisyys
  - Kapasiteettilähtöisyys (luodaan ymmärrys ja osaaminen hyödyntää AI:ta eettisesti)

- ”GDPR for dummies”

ETAIROS

Taustaa teemoihin ja käyttötapauksiin



Sosiaali- ja terveyspalvelut

SOTE-alan haaste

- Ristiriitaiset kehityskulut:
- Tieteellis-tekniset innovaatiot pidentävät elinaikaa ja lisäävät mahdollisuuksia hoitaa erilaisia vaivoja ja tiloja.
- Väestön ikääntyminen ja raihnastuminen heikentävät huoltosuhdetta.
- Kaikkea mitä voitaisiin tehdä, ei ole mahdollista toteuttaa.
- Toiminnan painopistettä on suunnattava ennaltaehkäisyyn, toimintakyvyn ylläpitämiseen ja sosiaalisten voimavarojen vahvistamiseen.

Tekoälyn sovellusalat sote-aloilla

Tekoäly ja terveydenhoito Suomessa –projektin mukaan

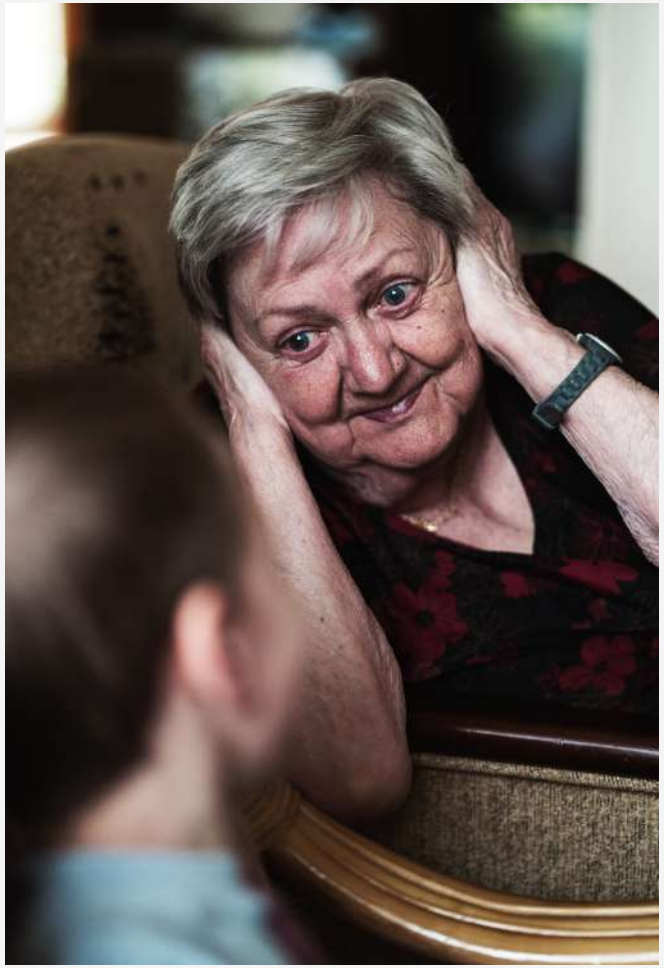
1. Kansallisen tason tiedonhallinta-alustan käyttöönotto
2. Uuden sukupolven potilastietojärjestelmien kehittäminen ja käyttöönotto
3. Tekoälyn ja läpimurtoteknologioiden yhteisvaikutuksen hyödyntäminen
4. Johtamiskäytännöt
5. Palveluprosessit
6. Asiakassuunnitelma
7. Ennaltaehkäisy ja omahoito
8. Tekoälyn hyödyntämiseen diagnostiikassa,
9. Lääkehuolto
10. Logistiikka

(Neittaanmäki & co Tekoäly ja terveydenhoito Suomessa 2019)

ETAIROS

- Erityisen kiinnostuksen kohteena
  - Yhteistoimijuutta tuottavat tekoälyjärjestelmät, joiden tavoitteena on
  - Asiakkaiden/potilaiden omaa aktiivisuutta hyödyntävä ennaltaehkäisevä ja terveyttä ja hyvinvointia ylläpitävä toiminta
- Tavoitteena
  - Luoda yhteistoimijuutta tuottava tekoälymalli, jonka varassa voidaan toteuttaa eettisesti kestäviä, tekoälypohjaisia sote-sovelluksia.

Ikääntyminen



ETAIROS

Kotihoidon taustaa

- Hyvinvointi- ja seurantateknologiavalikoima on laaja ja sekava.
- SOTE-palvelujen kustannusten kasvun hillitsemiseksi ja palvelujen laadun parantamiseksi palvelujen tuottaminen kotiin entistä laajemmin välttämätöntä.
- Hyvinvointiteknologian tehokas hyödyntäminen edellyttää yhteisiä **taustapalveluja** (esim. 24/7 päivystys, tekninen tuki, asiakasohjaus) ja **järjestelmäpalveluja** (esim. tietovaranto, tunnistautuminen, tietojen rakenteet).
- Kaivataan **palvelujärjestelmää** tukemaan ratkaisuja, jotka mahdollistavat tarpeiden ja palvelujen kustannustehokkaan ja eettisen yhteensovittamisen.

Tekoälyn sovellusalueet

Passiivinen ja aktiivinen teknologia

- Onnettomuuksien etämonitorointi ja hälytykset
- Paikannus
- Hyvinvointitieto
- Lääkehoito

Älykotikonseptit

- Big data -Toimintakyvyn ja aktiivisuuden seuranta proaktiivisesti
- Kerätyn datan analytiikka ja seuranta lapsenkengissä

Kotihoidon toiminnanohjausjärjestelmät

- AI:n hyödyntäminen?

Etairoksen kiinnostuksen kohteet

AI:n käyttötapaukset ja eettiset kysymykset kotihoidossa

- Asiakas / omainen / hoitohenkilökunta
- Esim. paikkatieto & alueelliset hyvinvointierot

Eettisyys AI-ekosysteemissä

- Toimijaverkoston tutkimus
- Tekoälyn eettisen hyödyntämisen ohjauskeinot

Tiedon omistajuus

- Omistajuuden eri vaiheet; tiedon säilytys / siirto / kaytto

Tiedolla johtaminen

- Tekoäly päätöksenteossa; moraalinen AI

Kansallinen turvallisuus

Kv. suhteet ja näkökulmat

Kv. asioihin kuuluvat esimerkiksi ulkoministeriön ja puolustusministeriön kehityskohteet

- Älykkäät teknologiat muuttavat toimintamalleja
- Syntyy uusia ongelmia ja haasteita toimintaprosessien muuttuessa
- Eettisten näkökohtien mukaan ottaminen inhimillisiä toimintamalleja suunniteltaessa
- Älykkäiden teknologioiden käyttömahdollisuudet

Yleisiä kehityskohteita

1. Säädösten ja tulevaisuuden kohtaaminen
2. Kompleksisuuden kasvu ja riskien lisääntyminen
3. Suunnittelija- ja toimittajavastuu
4. Sosiaalinen resilienssi ja muuttuvan yhteiskunnan etiikka
5. Etiikasta säädöksiin ja lakeihin
6. Etiikka politiikka ja juridiikka
7. Itsekorjaavuus ja riskien ennustaminen

Suunnitteluprosessi

- Formaalit ja sisällölliset välineet
- Elämäkeskeinen suunnittelu
- Prosessien ontologinen kuvaaminen
- Sosiaaliset implementointiprosessit

Yhteiskunnallinen vakaus

Tausta: Paljon dataa

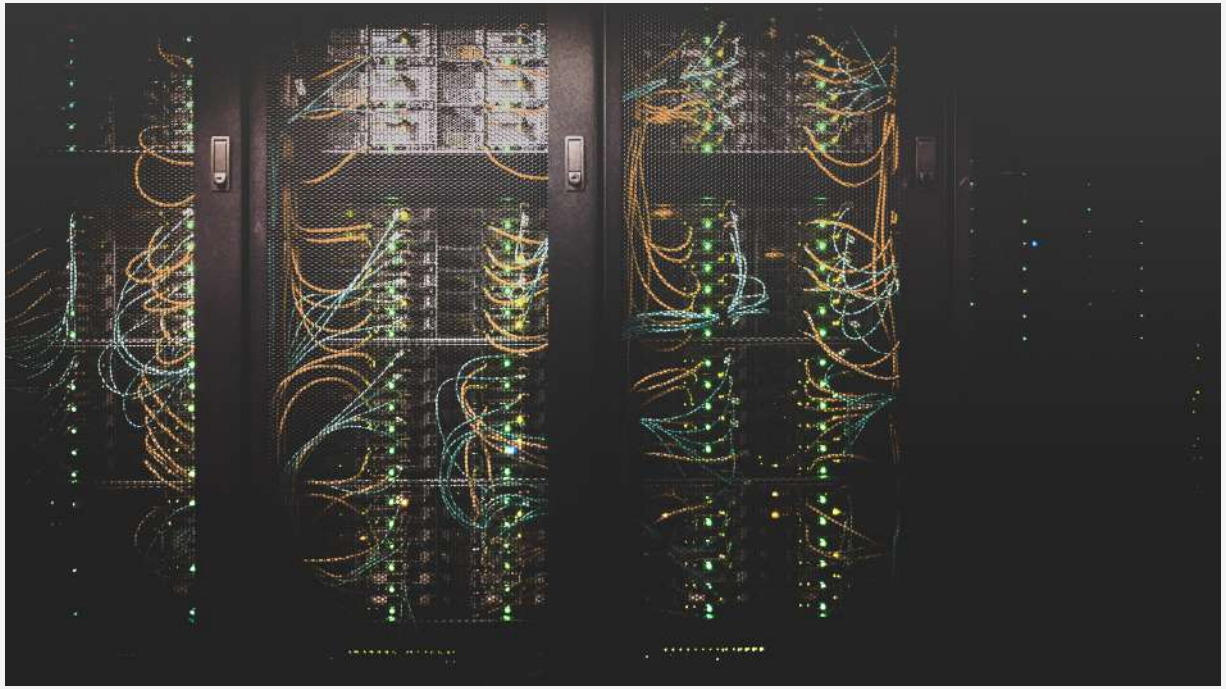
- Rakennusautomaation antureista / järjestelmästä (energiamittaus tähän kategoriaan)
- Paloilmoitinjärjestelmistä
- Rikos- ja kulunvalvontajärjestelmistä
- Kameravalvonnasta
- Pysäköintijärjestelmistä (sekä ilmaisin- että kamerat)
- Oviautomatiikasta
- Hisseistä
- Kaupallisten tietojen järjestelmistä teoreettisella tasolla
- Huoltokirjasta
- Palvelutuottajien raporteista
- Kiinteistön ja sen palvelutuottajien rfid-sovellutuksista

Tekoäly turvallisuushenkilöstön älykkäänä parterina

- Ihmisen on vaikeaa ottaa huomioon kaikki saatavilla oleva tieto tilannetta arvioitaessa
- Useimmat tilanteet lieviä vs väärä hälytys
- Tilanteet liittyvät kiinteistön hallintaan, henkilöturvallisuuteen, materiaalisten vahinkojen välttämiseen, jne
- Kun tilanteet tulevat ja menevät, järjestelmän sopisi oppia niistä
- Tarvetta sekä sääntöpohjaisuudelle että datapohjaisuudelle

Eettiset kysymykset ja ratkaisut

- Eettisiä kysymyksiä esim ongelman tunnistamisessa, tilanteiden profiloinnissa, reagoinnin tasossa (ei yli- eikä alireagointia)
- Eettiset ongelmat voivat liittyä järjestelmän hyödyntämiseen mutta järjestelmä voisi myös tunnistaa tilanteita joissa ollaan mahdollisesti eettisesti hankalan päätöksenteon edessä
- Tavoite: Mallinnetaan (spesifioidaan) yleisiä ratkaisuita sisältävää järjestelmää ja testataan sitä mallin kautta.
  - Haastatellaan henkilöstöä
  - Tarkastellaan suunnitteluprosessia: menetelmät, roolitukset, jne.



Autonomiset järjestelmät

ETAIROS

Autonomiset järjestelmät taustaa

- Autonomisessa järjestelmässä laitteistot ja ohjelmistot toimivat yhdessä ongelman ratkaisemiseksi. Jotta järjestelmää voidaan pitää itsenäisenä, järjestelmän on kyettävä keräämään tietoja, löytämään tietoon perustuva ratkaisu ja suorittamaan toimenpide tavoitteen saavuttamiseksi.
- Autonomisia järjestelmiä kehitetään autonomian astetta vähitellen nostaen.
- Tekoälyratkaisut mahdollistavat ihmisten ja osittain autonomisten järjestelmien välisen yhteistyön.
- Autonomiset järjestelmät liikenteessä:
    - Muun muassa merenkulku, miehittämättömät järjestelmät (ilma, vesi, maa, avaruus), kaupunkiliikenne, kaupunkien välinen ja maaseutuliikenne, työkoneet (mm. maa-, metsä- ja kaivostalous)

Tekoälyn sovellusalueita autonomisissa järjestelmissä

- Logistiset ratkaisut, esim. aukottomat logistiikkaketjut
- Älykäs energian käyttö
- Robottiautot
- Miehittämättömät laivat ja satamat
- Liikenne palveluna (Maas)
- Tieliikenteen seurantajärjestelmät
- Autonomiset työkoneet
- Teollisuusjärjestelmät

Kuva: RAAS

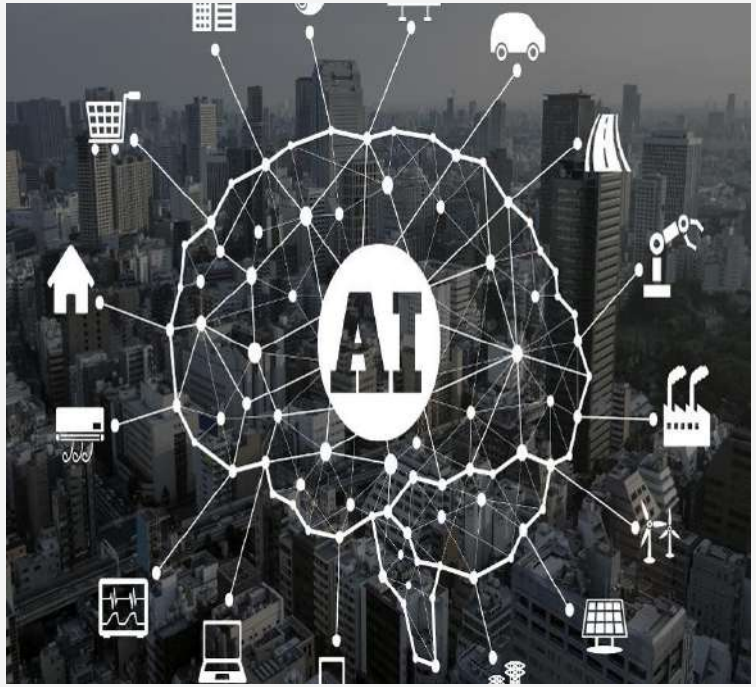
Etairoksen kiinnostuksen kohteet

- Tekoälyn ja ihmisen vuorovaikutus ja sen vaikutus esim. turvallisuuteen ja päätöksentekoon
- Autonomisten autojen integrointi liikennejärjestelmiin
- Tekoälyn etiikka paikannustiedoissa
- Kilpailuedun luominen eettisen tekoälyn avulla
- Ihmistenvälisen dynamiikan muutos tekoälyn ohjatessa eri tehtäviä
- Tekoäly ’työkaverina’ ja mentaalinen hyvinvointi
- Liikenteen autonomisten järjestelmien käyttöönoton pitkäjänteiset vaikutukset yksilöiden, organisaatioiden ja yhteiskunnan toimintaan

Ohjaus ja sääntely

Tekoälyn hyödyntäminen yhteiskunnallisessa päätöksenteossa ja ohjauksessa

- Hallinnossa on otettu asteittain ja lisääntyvästi käyttöön tekoälyä muun muassa päätöksenteon tukena, asiakaspalvelun automatisoituna tukena sekä tietojärjestelmien hyödyntämisessä.
- Nostaa esiin kysymyksen miten tekoälyn hyödyntäminen toteutetaan eettisesti ja hyvää hallintotapaa noudattaen eri näkökulmat huomioiden (esim. luomatta tarpeettomia rajoitteita toiminnalle, mutta samalla kansalaisten yksityisyys suojaten).



Tekoälyn soveltaminen

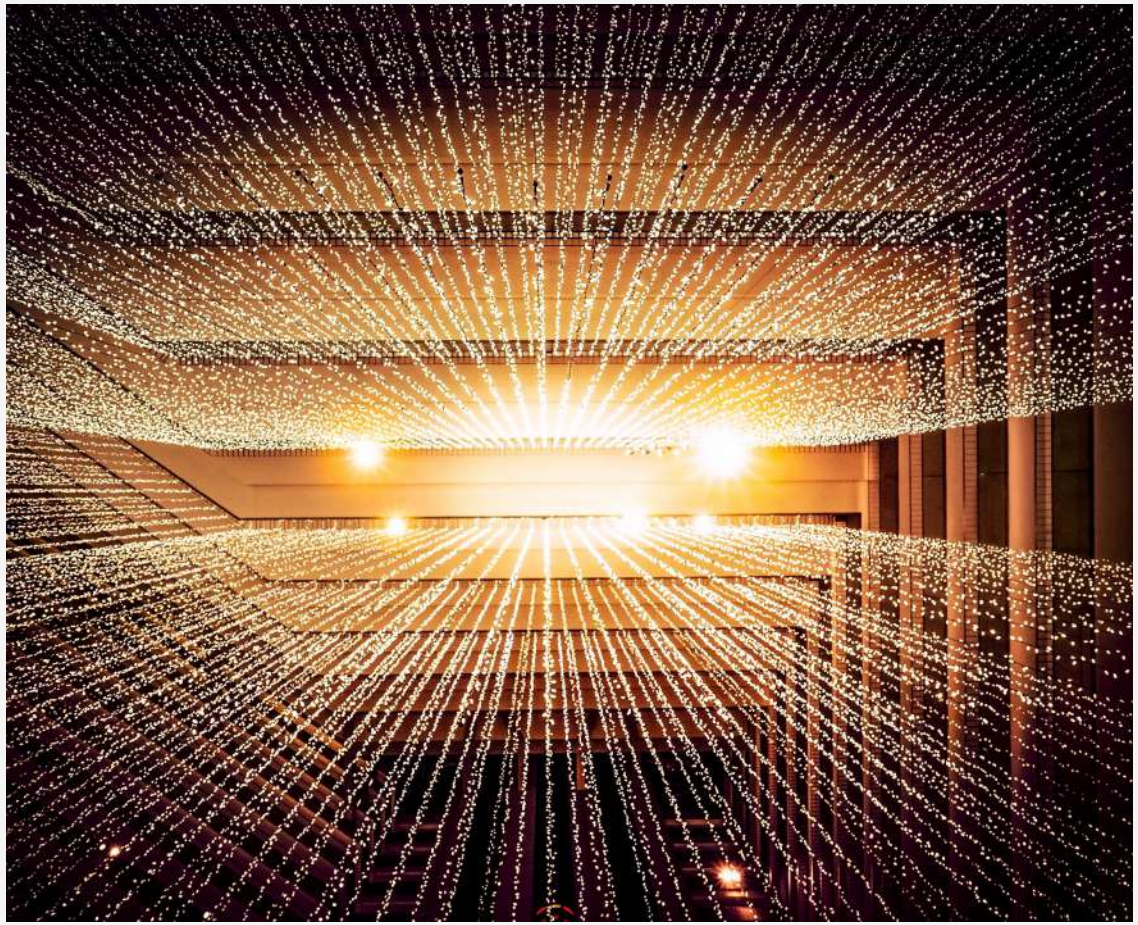
On ennakoitavissa, että julkinen palvelu ja viranomaistoiminta muuttuvat usein tavoin, esim. (Koivisto & Leikas ym. 2019)

- Tarkistusprosessi ja päätösten perustelujen esittäminen muuttuvat
- Yksilön huomiointi päätöksenteossa muuttuu
- Asiakkaan suhtautuminen viranomaiseen muuttuu
- Objektiivisuuskäsitys voi muuttua
- Työn luonne muuttuu
- Vaaditaan suurta vastenopeutta niin ratkaisujen tekemisessä kuin järjestelmien ylläpidossa
- Tietojen yhteiskäyttö lisääntyy

Etairoksen kiinnostuksen kohteita

- Eettisten normien tarve ja perustelut julkisessa hallinnossa ja palveluntuotannossa, toiminnan sertifikaatit ja läpinäkyvyys
- AI:n käyttö julkisten palveluiden proaktiivisessa tarjonnassa
- AI kuntien ja kaupunkien käytössä
- Datan varastointi ja rekisteriviranomaisen vastuu
- Eettisten ohjeiden poliittiset ja juridiset muutokset
- GDPR:n rooli
- Itsesäätely vs. julkinen regulaatio
- Julkiset hankinnat

Päätössanat



Päätössanat ja hankkeen seuraavat askeleet

Seuraavat askeleet

- Teema-alueiden työn käynnistyminen ja sisältöjen tarkentaminen
- Yhteydenpito kumppaneihin
- Seuraavat suuremmat tapahtumat
  - Avoin dialogifoorumi alkuvuodesta 2020 ja
  - Seuraava yhteiskehittämisen foorumi keväällä 2020

ETAIROS
Kiitos osallistumisestasi!
Yhteyshenkilöt ja lisätietoja: www.etairos.fi