

Eettisen suunnittelun diskurssi: Älykkään teknologian ja sen etiikan suunnittelu - Esimerkkitapaus diabetes



Ohjeistus dokumentin käyttöön

Tämä dokumentti sisältää eettisen suunnitteludiskurssin kuvauksen, jossa tapaus esimerkkinä toimii II-typin diabeteksen ehkäiseminen. Ontologia, jolle tapaus esimerkki perustuu, on luonteeltaan yleismaailmallinen. Toisin sanoen sitä voidaan hyödyntää minkä tahansa suunnitteludiskurssin ohjaamiseksi.

Ontologian peruselementit:

- Toimintaperheet
- Toimijat
- Toiminnan kohde
- Kontekstit
- Teknologiat ja tekniset ratkaisut
- Suunnittelun rajapinnat.

Dokumentin sivut 3–16 käsittelevät tapaus esimerkkiä.

Sivut 17–19 käsittelevät ohjeistuksia ja mallin diskurssin jäsentämiseksi.

Dokumentti on tarkoitettu selailtavaksi, eli sen sisältöä tarkastelemalla voi reflektoida, onko omassa suunnittelutyössä jokin keskeinen näkökulma jäänyt vaille huomiota, kun tarkoitus on istuttaa eettisen diskurssin näkökulmia osaksi suunnittelutyötä.

Olemme aiemmin julkaisseet dokumentin, jossa keskustellaan yleisesti älykkäiden teknologioiden suunnittelun ja etiikan suhteesta: <https://etairos.fi/wp-content/uploads/2022/03/etairos-etiikka-ja-alykkaiden-teknologioiden-suunnittelu.pdf>



1. EETTINEN suunnitteludiskurssi



Toimintaperheet

Toimintaperheen kuvaamien toimintojen analyysin rakenne:

Toiminta ja sen analyysi on lähtökohta:

- toiminnan kuvaus
- toiminnot ja liitännäistoiminnot
- alitoiminnot.

Teknologia ← ihminen & tekniikka.

Teknologian rooli toiminnassa.

Tavoite. Mihin suunniteltavaa tekniikkaa käytetään.

Toiminta voi liittyä niin poliitikon toimintaan, sairaalaan hallinnon toimintaan, lääkärin toimintaan jne. Toiminnot liittyvät visioon, mutta eivät välttämättä suoranaisesti toisiinsa.

Toimintaperheiden kuvaus diabetes-tapausesimerkissä:

Päätöksenteko ja suunnittelu / Hallinto

- strategiset investointipäätökset
- hyvinvointialueet ja sairaalataso.

Ehkäiseminen

- tiedottaminen
- koulutus
- työpaikan toimenpiteet
 - o työympäristö.

Hoitotoimenpiteet/Sairaudenhoito

- terveydentilan seuranta
 - o diagnoosi

- interventiot.

Tutkimustoiminta

- uuden tiedon tuottaminen
- ohjausvaikutukset.

Kansalaistoiminnot

- yksilön elämäntavat
 - o liikunta
 - o ruokavalio
 - o terveydentilan seuranta.

Ja niin edelleen.

Avoimia kysymyksiä:

Miten edetä siihen tilanteeseen, että asiat ovat toivotulla tavalla (nykytila vrs. tavoitetila)?

Mitä ”toteutamme oikeusvaltioperiaatetta” tarkoittaa käytännön suunnittelutyössä?

”Hyvepesu”?

Toimintaperheet, etiikka

Etiikka: mikä tekee kuvatusta toiminnasta suotavaa ja hyväksyttävää? Onko sivuvaikutuksia?

Reunaehdot:

Yhteiskunnan arvot, lainsäädäntö.
 Ihmisoikeudet.
 Kansalaisyhteiskunnan oikeudet.
 Taloudelliset, terveydelliset ja sivistykselliset oikeudet.
 Ryhmäoikeudet.
 Demokratia.
 Oikeusvaltioperiaate.

Viranomaiset.
 Hyvinvointiyhteiskunnan ihanne.
 Tasa-arvo.
 Kattava asiantuntijuus ja yhteistyö suunnittelussa.
 Muiden kunnioittamiseen perustuva sosiaalinen elämä.
 Luonnon kantokyky.

Yksittäisten toimeenpanijoiden taso:

Kokemus tavoitteiden mielekkyydestä.
 Tunnistaa, että toimet ovat mielekkäitä, perusteltuja ja mahdollisia
 – tällä hetkellä mahdollista ja ”mahdottoman” mahdolliseksi tekeminen.
 Koettu hyvinvointi ja tavoiteltu hyvinvointi
 – ensimmäisen persoonan ja kolmannen persoonan näkökulmat.
 Arkitietämys ja ammattilaitietämys.

Organisaatioiden taso:

Iso määrä eri toimijoita, jotka itseorganisoituvat.
 Perhe, yritys, poliittinen toiminta jne. eri tasojen organisaatiot.

Avoimia kysymyksiä:

Stasis ja dynamis (joustamattomuus ja pysähtyneisyys verrattuna joustavuus ja toiminnallinen)?
 Etiikan rajat? Prudentiaalisuuden ja eettisyyden välinen ero?

Toiminta

Toiminta: Hallinto

Mahdolliset hallinnolliset toimenpiteet

Kansallinen päätöksenteko.
Kunnallinen päätöksenteko.
Hyvinvointialueen päätöksenteko.
Sairaushoito organisaatio.
Koulutuslaitokset.

Toiminta: Hallinto, etiikka

Etiikka: mikä tekee kuvatusista toiminnasta suotavaa ja hyväksyttävää? Onko sivuvaikutuksia?

Viisi hyvän hallinnon periaatetta:

- yhdenvertaisuus
- tarkoitussidonnaisuus (vrt. oikeusvaltioperiaate)
- objektiviteetti (toiminnan perusteltavuus ja puolueettomuus)
- suhteellisuus (toimien suhde päämäärään)
- luottamuksensuoja.

Yhteiskunnan arvot, lainsäädäntö.
Ihmisoikeudet.
Kansalaisoikeudet.

Taloudelliset, terveydelliset ja sivistykselliset oikeudet.
Ryhmäoikeudet.
Demokratia.
Oikeusvaltioperiaate.
Viranomaiset.
Hyvinvointiyhteiskunnan ihanne.
Tasa-arvo.
Kattava asiantuntijuus ja yhteistyö suunnittelussa.
Muiden kunnioittamiseen perustuva sosiaalinen elämä.
Luonnon kantokyky.

Toimeenpano:

Miten korkeat ja abstraktit periaatteet saadaan kuljetetuksi suunnitteluun?

- Koulutus, omaehtoinen opiskelu, tarkistuslistat, esimerkit, herkistyminen toteuttamaan eettisesti korkeatasoista toimintaa.

Kohtaamisen taso

- esim. sosiaalisen kohtaamisen tarkistuslistat luovat luottamusta asiantuntijan pääomaan.
- empatiakyky, ystävällinen käyttäytyminen ja vuorovaikutus, esimerkin voima (esim. käyttäytyminen osa kulttuuria)
 - o esim. poliisien koulutus ”hankalan” asiakkaan kohtaamisesta, lääkäri kohtaa koko kansakunnan kirjon.

Toteuttavalla portaalla on riittävästi toimivaltaa ja resursseja kehittää omaa työtään ja pohtia sitä

- tiedollinen resurssi, aika, fasiliteetit, päätösvalta, jne.

Avoimia kysymyksiä:

Muutosagenttiuden luonne?

Toiminta: Ehkäiseminen

Yhteisötaso

- Tiedottaminen
 - tiedotusvälineet.
- Ympäristön suunnittelu.
- Normit
 - Lainsäädäntö.

Yksilötaso

- Liikunta.
- Ruokavalio.
- Stressinhallinta.

Toiminta: Ehkäiseminen, etiikka

Etiikka: mikä tekee kuvatusta toiminnasta suotavaa ja hyväksyttävää? Onko (ristiriitaisia sivuvaikutuksia), oheisvaikutuksia (muuttaa kulttuuria tavalla, mitä ei ole haluttu, muita seurauksia mitä ei ole tarkoitettu)? Minkä takia toiminta vaikuttaa perustellulta? Tavoitellut vaikutukset.

Yhteisötaso

- Ihmiskeskeinen ja ennakointikykyinen yhteiskunta.
- Riski-indikaattorit ja sosiaaliset tunnusmerkit.
- Data nähtävä kansallisena pääomana kaikkien elämän parantamiseen.
- Pahan korjaaminen ja estäminen, hyvän ennallistaminen, säilyttäminen, vahvistaminen ja luominen.
- Varautuminen ja edellytysten luonti
 - maailman väistämätön muuttuminen, dynamis ja stasis.
- Kesto (pitkäkestoisuus), resurssit.
- Yhdenvertaisuus.
- Hyvinvoiva – ei-hyvinvoiva? Terve – ei-terve? Halutaan ehkäistä, että menetetään hyvä
 - terveys on suhteessa edellytyksiin (esim. synnynnäinen geenivirhe tai rakennemuutos, suhde ikään jne.)
 - käytetään samoja termejä kuin esim. taloudessa, joka voi johtaa ajattelua harhaan.

- Eri asiantuntijoilla tai tahoilla voi olla erilaiset näkemykset ihmisten osallisuudesta ja sosiaalisesta roolista.
- Ideaali, johon nykytilaa verrataan (liukuma), yhteisöllinen ja yksilöllinen.
- Me-ryhmä yhteisen sosiaalisen toiminnan ympärillä.
- Terveiden varaan muodostuva yhteisö?
- Kulttuuri
 - riitit ja riittävyys (kulttuuriset rituaalit), joiden avulla hallitaan kulttuuria ja istutetaan erilaisia sisältöjä ja elementtejä terveyden edistämiseen?
- Sosiaaliset mallit mitä toivotaan, roolit.
- Sairaan ja terveen rooli? Ei jumittavia, vaan tavoitetta edistäviä
 - esim. länsimainen ja itämainen terveysajattelu: hoidetaan vai ehkäistään?

Yksilötaso

Yhteiskunnallinen eriarvoisuus.

Elintapojen säätely ja rajoitukset.

Riskitekijöiden poistaminen ja pahenemisen estäminen (reaktiivisuus), hyvän säilyttäminen.

Toimijuuden vahvistaminen, autonomia (toiminta tavoitteiden puolesta tai vastaan). Miten saadaan rooli Potilaasta aktiiviseen toimijaan?

Asiantuntijoiden ja vertaisten tuki.

Käytettävän tekoälyn luotettavuus: palvelee tarpeita yksilön parhaaksi.

Osallisuus.

Normaalitilan ja rutiinien muuttaminen, ulkoiset ja sisäiset resurssit: kenen ihanteet ja standardit?

Auttaminen: auttajan ja autettavan asema ja edellytykset ja niihin vaikuttaminen.

Autettavan / vaikutettavan osallistuminen ja oma panos, omien intressien, uhan, seurausten ja auttamistoimien tunnistaminen.

Vastavuoroinen vaikuttaminen, luottamus.

Edellytykset terveelliselle, turvalliselle ja luotettavalle toiminnalle.

Estää tilannetta pahentumasta, säilyttää kohteen hyvät.

Tarjoaa ulkoisia resursseja, luo ja vahvistaa sisäisiä resursseja.

Yleisluontoinen ja reaktiivinen apu.

Ideaali, johon nykytilaa verrataan (liukuma), yhteisöllinen ja yksilöllinen, josta yksilö voi tunnistaa oikean tai suotuisan suunnan.

Yksilöiden välillä kasvava ero (ennestään hyvinvoivat pääsee käsiksi uusiin mahdollisuuksiin, ennestään heikommassa asemassa olevat menettävät vanhatkin mahdollisuudet).

Oma sosiaalinen tausta vaikuttaa toimintaan (esim. matala koulutus ja halu kouluttautua).

(Addiktiivisen) energian siirtäminen positiiviseen toimintaan esim. riittien ja käytänteiden oppimisen myötä.

Sairaana ja terveen rooli? Ei jumittavia vaan tavoitetta edistäviä. Mahdollisuus vaihtaa identiteettejä. Mallit hahmottaa todellisuutta

- esim. (sairaana) ihmiset palaavat uudestaan ja uudestaan negatiivisiin ajatuksiin
- esim. sosiaalisten verkostojen merkitys yksilölle.

Pelko tietää piilevästä sairaudesta. Muutoksen logiikka ja muutoshaluttomuus.



Toiminta: Hoitotoimenpiteet

Interventiot.

Diagnostiikka

- algoritminen päätöksenteko.

Terveydentilan seuranta

- esim. anturit.

Opastus

- sokeritasapainon seuraaminen
- liikunta ja ruokavalio-ohjeistukset
- yksilön toimijuus ja vastuu (jalkauttaminen).

Lääkitys

- esim. insuliinipistokset, pillerit, insuliinipumput
- kevyt tai raskas hoito.

Ruokavalio.

Toiminta: hoitotoimenpiteet, etiikka

Etiikka: mikä tekee kuvatusta toiminnasta suotavaa ja hyväksyttävää? Onko sivuvaikutuksia?

Normaalitilan ja rutiinin muuttaminen, ulkoiset ja sisäiset resurssit

- tottumukset, riittävä valmius luopua haitallisista rutiineista, kannusteita ja resursseja luoda uusia rutiineja, samalla kriittinen ote rutiineihin ja halu muuttaa rutiineja
- taustasta huolimatta ihmisillä tulee olla mahdollisuus hypätä samaan kansanterveydelliseen kelkkaan
- kenen ihanteet ja standardit?

Auttaminen: auttajan ja autettavan asema ja edellytykset ja niihin vaikuttaminen.

Autettavan / vaikutettavan osallistuminen ja oma panos, omien intressien, uhan, seurausten, auttamistoimien, elämän suunnan tunnistaminen.

Vastavuoroinen vaikuttaminen, luottamus.

Edellytykset terveelliselle, turvalliselle ja luotettavalle toiminnalle.

Estää tilannetta pahentumasta, säilyttää kohteen hyvät.

Tarjoaa ulkoisia resursseja, luo ja vahvistaa sisäisiä resursseja.

Yleisluontoinen ja reaktiivinen apu.

Elintapojen ja rutiinien muutos: vahinkojen estäminen ja korjaaminen, hyvän suojelu, uuden hyvän luominen.

Vertaistuessa kaikki samassa tilanteessa, voidaan löytää tasapaino itsensä kanssa (esim. AA-kerho).

Sairaana ja terveen rooli? Ei jumittavia vaan tavoitetta edistäviä.

Potilaan tila nyt ja tulevaisuudessa

- vaikuttaa esim. millaisia hoitotoimenpiteitä tai lääkityksen vahvuuksia on.

Pelko tietää piilevästä sairaudesta, muutoshaluttomuus.

Avoimia kysymyksiä:

Millaista teknologiaa pystytään kehittämään eri tilanteisiin?

Toiminta: Tutkimus

Perustutkimus

- lääketieteellinen
- käyttäytymisen ohjaaminen
- keinot.

Sovellusratkaisut

- algoritmit ja tekoäly.

Teknologiaturkimus

- tekniikka.

Sosiaalitutkimus / ihmistutkimus

- indikaattorit datan perusteella
- käyttäjään liittyvät toimenpiteet.

Tutkimuksen etiikka

- tutkimuksesta tiedottaminen
- tutkimuksen oikeudet
- intressiristiriidat (esim. tutkimuksen ja kaupallisen välillä)
- tietoturvakysymykset
 - o sensitiivinen data
 - o ei-sensitiivisestä datasta sensitiivistä? (esim. vain yksi henkilö maassa, jolla on tietty sairaus → anonymisointi ei riitä).

Toiminta: Tutkimus, etiikka

Etiikka: mikä tekee kuvatusta toiminnasta suotavaa ja hyväksyttävää? Onko sivuvaikutuksia?

Tutkimuksen merkityksellisyys ja hyödyllisyys (perustelut).

Tutkijoiden koskemattomuus.

Hyödyt ja taakat on oikeudenmukaisesti jaettu.

Osallistujien riskit on minimoitu ja oikeutettu hyödyn perusteella.

Osallistujien kunnioitus.

Eettiset ohjeet ja lainsäädäntö.

Tietoinen suostumus: tutkittava informoidaan tutkimuksesta yksityiskohtaisesti, hän ymmärtää sen riittävällä tasolla, ilmaisee suostumuksen, sallii intervention.

Potilaan tietolomake.

Luottamuksellisuus: tieto saatavilla vain valtuutetuille tahoille, vain tarpeellisen datan keräys, tallennus, suojaus

Yksityisyys: tutkimuskohde päättää.

Kahdenkeskinen kommunikaatio (*privileged communication*): osapuolten välinen, luottamuksellinen, turvallinen ja luotettava.

Arvostus ja vastuu: ihmiset, totuus ja tutkimus.

Kuka saa käyttää dataa ja miten? Terveys- ja sosiaalidata kansallisomaisuutena. Käytetty versus alikäytetty.

Tulokset ymmärrettäväksi kansalaisille?

Viestin tyyli, mahdollisuuksien tarjoaminen.

Kansalaisten erilaiset selviytymismekanismit, eri elämäntilanteet, tunteet ja asioiden merkitykset itselle.

Avoimia kysymyksiä:

Onko tutkijalla oikeus vetäytyä tutkimuksesta? Julkisesti jaetun datan riski- / hyötyanalyysi? Hallinnon rakenteet ja ohjeet teknologian käyttöön? Tutkimuseettinen koulutus tutkijoille?

Toiminta: Kansalaistoiminnot

Yksilön elämäntavat

- liikunta
- ruokavalio
- terveydentilan seuranta.

Ryhmätoiminnot.

Tiedottaminen

- sosiaalinen media
- lehdistö

- kohdennettu
- asiantuntijaorganisaatiot (esim. apteekki, lääkärit, työterveys, neuvola...)
- vertaiset (esim. martat, somen "osaamispääoman" yhteisöt, influencerit / vaikuttajat).

Kansalaisjärjestöt.

Toiminta: Kansalaistoiminnot, etiikka

Etiikka: mikä tekee kuvatusista toiminnasta suotavaa ja hyväksyttävää? Onko sivuvaikutuksia?

Kommunikaatio, diskurssi versus tiedottaminen, diskurssin ohjaaminen.

Influencerin "elämän" kesto.

Institutionalisointi, irtaantuu alkuperäisestä tarkoituksesta, politisoituminen

- päätökset eivät vastaa sitä mihin kansanterveydellisesti rahaa tarvittaisiin
- kansalaismoraali, "kaikkea kunnollista sopivasti".

Miten pystytään toimintamalleja muuttamaan, kun yhteiskunta muuttuu?

Institutionaalinen sitoutuminen konseptina vanhentunut tai vieras, tarvittaisiin uusia toimintamalleja tai vanhojen toimintojen modernisointia.



Toimijat

Toimijoiden osaaminen, tietämys.
 Arvot ja arvopohjainen suunnittelu.
 Roolit (potilas, lääkäri, riskiryhmäläinen, hallinto, tutkimus).
 Mentaaliset mallit, asenteet, persoonallisuus.
 Motiivit ja intressit.
 Sosiaaliset toimintamallit

- suhde yksilön toimintaan.

Kansalainen.
 Läheiset.
 Hoitohenkilökunta

- hoitajat
- lääkärit.

Poliitikko.
 Virkahenkilökunta

- hyvinvointialueet.

Sairaalahallinto.
 Työyhteisö.
 Farmaseutit (apteekin henkilökunta).
 Tiedottajat.
 Oppilaitokset.
 Datalähteet (esim. armeija, neuvola).

Toimijat, etiikka

Etiikka: mikä tekee kuvatusista toiminnasta suotavaa ja hyväksyttävää? Mikä on toimijoiden moraalit (teoria ja arki)?

Vastuullisuus.
 Eri osaamisalueiden asiantuntijat ja sidosryhmät, inklusiivisuus.
 Arvot, ihanteet ja periaatteet.
 Kokemukset, muistot, tunteet, arvot, identiteetti, kieli, käytännöt, osallisuus yhteiskuntaan ja sosiokulttuuriseen maailmaan.
 Käyttökulttuuri.
 Hyvinvointi, universaali käytettävyys, autonomia, yksityisyys, luottamus, vastuut.
 Etiikka riippuu toimijasta: toimijayhteisö (esim. sairaanhoidon etiikka), ala / yksilökohtainen (hoitajaetiikka), ydintoiminta.
 Otettava huomioon kokonaisvaikutukset muihin toimijoihin

- toimijayhteisön hallinnollinen diskurssi.

Käypähoito – kliininen autonomia – potilaan ja asiakkaan itsemääräämisoikeus – mahdollisuus vaihtoehtoiseen ajatteluun

- kommunikoinnin etiikka vaarallisista asioista potilaille. toimenpiteillä on hyötysuhde, yhteinen ja yleinen hyöty, yksilön oma tilanne
- potilaalla oikea ymmärrys lääketieteellisen tiedon luonteesta (taustatiedot, illuusiot jne.)
- lääkärin ammattitaito ja ammatillinen ymmärrys
 - o taustalla valtava määrä aiempaa kokemusta ja oppimista tapauksista (ekspertiisi ja praksi).

Miten hoitoyhteiskunnan diskurssi pitäisi käydä? Kartoitettava yleisiä ja roolikohtaisia eettisiä kysymyksiä, kehitettävä ymmärrystä

- lääketiedettä kansalaisille
- potilaan oman elämänsä asiantuntijuus (kommunikaatio, mitä ihmisten tulisi tietää, asioiden mieltäminen vaatii aikaa)
- asiantuntijuuden etiikka (esim. lääkärit, lääkärin kommunikaatiotaidot, esim. asioiden liioittelu potilaalle, ”kaksoisymmärrys” eli lääkärin omakohtaiset kokemukset sairauksista ja hoidosta).

Toiminnan kohde

Ympäristön osa, johon sovelluksella vaikutetaan.
Ihmistoimintojen muutokset ja niiden edistäminen.

Yksilöt, ryhmät ja organisaatiot.

2-tyypin diabeteksen ehkäiseminen ja hoito

- taudin puhkeamisen ehkäiseminen / lykkääminen
- taudin pahenemisen estäminen
- tiedottaminen.

Riskiryhmään kuulumisen selvittäminen

- mitä toimintoja kohdentuu, jos kuuluu riskiryhmään
- mitä yksilön itsensä tulee tehdä, jos kuuluu riskiryhmään.

Elämäntoimintojen muutokset ja niiden edistäminen

- käyttäytyminen
- kunnan kohentaminen
- sokeritasapaino
- ruokailutottumukset.

Ehkäisevien resurssien parantaminen

- hoitoresurssien takaaminen
- ympäristöresurssien takaaminen
- tukipalveluiden takaaminen.

Hoitoyhteisöjen toimintamallit.

Toiminnan kohde, etiikka

Etiikka: mikä tekee kuvatusista tavoitteesta suotavaa ja hyväksyttävää? Onko sivuvaikutuksia?

Yksilön näkökulma:

Käytettävyys ja yhdenvertaisuus

- myönteinen käyttäjäkokemus
- käyttäjätarpeet ja niiden liittäminen viranomaistoimintaan.

Kokemus tulevansa otetuksi vakavasti.

Asioiden (hyvien) laittaminen järjestykseen (välittömät, pitkäkestoiset, näkymättömät yms. hyödyt ja vaivat).

Organisaatio:

Toimintamallien ja -kulttuurin uudistamisen suunnittelu (tehtävänjako, osasto- ja viestintärakenteet, valta-, vastuu- ja kontrollisuhteet).

Muutoksen mahdollisuudet ja riskit.

Kokee tekevänsä jotain mielekästä.

Työkulttuuri, miten kehitetään.

Yhteiskunnallinen keskustelu kuolemista ja kuolemasta (toimenpiteet ennen kuolemaa, epävarmuus toimenpiteiden vaikutuksesta).

Taitojen häviäminen.

Teknisten muutosten varmistamisen etiikka (lääkärit eivät osaa enää toimia ilman tekoälyä).

Kontekstit

Toimintaympäristön osa, joka vaikuttaa toimintaan, mutta ei ole sen kohteena

Sosiaalinen

- Vertaiset.
- Yhteisö(t)
 - o normit ja asenteet.
- Käytännöt, säännöt ja lait.
- Sosiaalisen kanssakäymisen muutos.
- Yksilön terveys on sosiaalinen asia
 - o itsensä kanssa samankaltaisten ryhmien löytäminen
 - o pikkuidentiteettien rakentaminen.

Yksilöllinen

- yksilöiden persoonallisuus ja temperamentti
- tavoite
- tavat ja toimintamallit
- biologinen
 - o sokeritasapaino.

Mentaalinen ympäristö

- stressi, sosiaaliset paineet
- työpaikalla
 - o esim. kokemus taukojen mahdollisuudesta

- vaatimusten muutos

Fyysinen ympäristö

- maantiede
- rakennettu ympäristö
- luonnon ympäristö
 - o sää
- koti
- työpaikka
 - o diabetesergonomia
 - arkiliikunta
 - ruoka & välipalat.

Elämäntapa

- työrituaalit
- organisaatio
- työttömyys
- eläkeläisyys
- elintaso, varallisuus
- koulutus
- jne.

Kontekstit, etiikka

Käyttökontekstia vastaavien tarpeiden tarkastelu moraaliperiaatteiden varassa

Tulevaisuus ja toimintaympäristöjen muutokset

- megatrendit
- hiljaiset signaalit
- tekoälyn kehittymisen trendit.

Tulevaisuuden suuntaaminen: Strategiat, skenaariot, tiekartat ja tulevaisuuspolut.

Tulevaisuuden arvot ja etiikka: utopiat ja dystopiat, tekoälyn soveltaminen.

Yksilön terveys on sosiaalinen asia

- itsensä kanssa samankaltaisten ryhmien löytäminen
- pikkuidentiteettien rakentaminen.

Yhteisön voima

- miten rakennetaan yhteisöllisiä kertomuksia, joihin yksilö voi liittyä
- suurista kertomuksista moninaiseen valikoimaan erilaisia pienempiä kertomuksia (*liquid modernity*).

Erilaiset diskurssin tasot

- toden ja valheen ero
- arvot ja normit, joihin halutaan linkittyä (miten halutaan tehdä jokin asia toivottavan näköiseksi, joka vaikuttaa siihen, miten yksilö / ryhmä siihen reagoi).

Teknologia ja tekniset ratkaisut

Teknologia ohjaa, tiedottaa, opastaa, tukee ja hoitaa.

Tekniikat ja applikaatiot.

Tekoälytekniikalla on asetettavissa eettisiä toimintarajoitteita.

Tutkimusanalytiikka.

Tekoälytukijärjestelmät

- malli toiminnoista
- simulaatio päätöksenteon seurauksista
- yksilön seuranta- ja sensoritekniikat.

Käytännön sovellukset

- älypuhelimien liikuntasovellus
- valmentava chatbot.

Teknologia ja tekniset ratkaisut, etiikka

Palvelee suunniteltua tarkoitusta

- toivotut ja ei-toivotut vaikutukset, laajemmat elämänmuotoon vaikuttavat vaikutukset.

Järjestelmän asettamat roolit, viemät resurssit ja antamat konkreettiset hyödyt.

Vahingon välttämisen periaate, elämänlaadun parantaminen (ihmiset, ympäristö, yhteiskunta).

Sujuvuus ja helppokäyttöisyys, miellyttävyys ja motivoivuus.

On osa ihmisen toimintaa ja integroituu siihen.

Vastaa muutokseen ja korjaa puutteita, kehittää parempia toimintamalleja.

Huomioi tavoitteena olevan valmiin järjestelmän ominaisuudet ja käyttöönoton toimintaa uudistamalla.

Suunnittelu: yhteiskehittäminen, käytettävyys, ohjelmistotekniikka, osallistava ja arvosensitiivinen suunnittelu, hankinta, testaus.

Eettinen analyysi: mitä, miksi, vaikutukset.

Innovaatiot: hyödyllisyys ja mielekkyys käyttäjälle.

Käyttöönotto, käyttäjädatta, koulutus ja osaaminen, käyttökulttuuri, viestintä, elinkaari, jatkuva arviointi, edelleen kehitys.

Hyvintoimimisen kriteerit: toiminnan suunta, päivitys, muokattavuus, poistaminen.

Avoimen yhteiskunnan diskurssi

- eri näkökantojen tuominen esille, ongelmien ja niiden ratkaisujen ideointi ja tarkastelu, kyky nähdä toisten ongelmien (teesi, antiteesi, synteesi)
- ei-tietämisen tila, toiminta perustuu sen hetkiseen parhaaseen tietoon.

Uusien ratkaisujen käyttöönotto voi muuttaa käyttäjän identiteettiä, ajattelutapoja ja toimintatapoja

- mitä menetetään ja mitä saadaan tilalle? esim. pelkkä ”parempi terveys” ei riitä
- mielekkyys ja tuki.

Teknisten ratkaisujen eettisyys yritysten arjen suunnitteluprosesseissa

- halutaan ja osataan ottaa huomioon.

Suunnittelun rajapinnat

Human digital twin

- informaation prosessointimalli
- Ideal – Exception – Correction -malli (IEC)

HPI-rajapintojen määrittely (requirements).
(BDI, Beliefs-Desires-Intentions).

GOMS, ACT-R

- computational rationality (ACT-R, Bounded rationality & reinforced learning, jossa toimintaohjeiden valinta / tuottaminen on automatisoitu).

Requirements engineering.

Suunnittelun rajapinnat, etiikka

Palvelee suunniteltua tarkoitusta.

Järjestelmän asettamat roolit, viemät resurssit ja antamat konkreettiset hyödyt.

Vahingon välttämisen periaate, elämänlaadun parantaminen (ihmiset, ympäristö, yhteiskunta).

Sujuvuus ja helppokäyttöisyys, miellyttävyyys ja motivoivuus.

On osa ihmisen toimintaa ja integroituu siihen.

Vastaa muutoksiin ja korjaa puutteita, kehittää parempia toimintamalleja.

Huomioi tavoitteena olevan valmiin järjestelmän ominaisuudet ja käyttöönoton toimintaa uudistamalla.

Suunnittelu: yhteiskehittäminen, käytettävyys, ohjelmistotekniikka, osallistava ja arvosensitiivinen suunnittelu, hankinta, testaus.

Eettinen analyysi: mitä, miksi, vaikutukset.

Innovaatiot: hyödyllisyys ja mielekkyys käyttäjälle.

Käyttöönotto, käyttäjädata, koulutus ja osaaminen, käyttökulttuuri, viestintä, elinkaari, jatkuva arviointi, edelleen kehitys.

Hyvintoimimisen kriteerit: toiminnan suunta, päivitys, muokattavuus, poistaminen.

Kuinka määritellään hyvät päämäärät, toimintatavat, jotka eivät riko eettisiä periaatteita (vahingollisuus)

- suhteessa päämääriin
- suhteessa muihin keskeisiin hyviin.

Tekniikoilla kaksoiskäyttö (hyvään ja pahaan): älykkääseen teknologiaan voidaan rakentaa esteet pahalle

- kasvattamalla kykyä suurempaan hyvään kasvatetaan kykyä myös suurempaan pahaan.

Millaisella diskurssilla ratkaistaan suunnittelun rajapinnan ongelmat

- esim. suunnittelijoiden omat implisiittiset oletukset ja ymmärrys erilaisista käsitteistä.

Vapaudet ja rajoitukset

- seurausvaikutusten ottaminen huomioon
- arkielämän diskurssit: millä tavalla rajoitteet perustellaan, rajoitteiden mielekkyys.

2. Työkalu



Yleistä

Society 5.0 (S5.0) ajattelu on tekoälysuunnittelun perustana.

Kokonaisvaltainen ajattelu – tekniikka ihmisen toiminnan ja elämänlaadun parantajana.

Suunnittelu on suunnittelukysymysten asettamista ja niihin vastaamista.

Määritellään peruskysymykset (ontologia).

Kysymysten periytyminen.

Ohjeet

1. Kirjaa visio.
2. Määrittele toimintaperhe / -kokonaisuus (vision kannalta relevantit toiminnat, yleisemmällä tasolla).
3. Analysoi toiminnat & eettiset kysymykset.
4. Analysoi kullekin toiminnalle toimijat & eettiset kysymykset.
5. Analysoi kunkin toiminnan tavoitteet & eettiset kysymykset.
6. Analysoi kunkin toiminnan kontekstit & eettiset kysymykset.
7. Analysoi teknologia ja tekniset ratkaisut & eettiset kysymykset.
8. Tarkenna toimintakuvaukset syventyen yksittäiseen toimintaan (yhdistä kohdat 2–7) sekä kuhunkin liittyvät eettiset kysymykset.
9. Visualisoi toiminta: tavoite, toiminta, toimijat, teknologia, kontekstit, prosessi (toiminnot & vuorovaikutukset), eettiset kysymykset, ongelmakohdat (riskit, puutteet, redundanssi, pullonkaulat, yms.), mahdolliset muut huomioitavat tekijät (esim. kustannukset).
10. Priorisoi ongelmakohdat ja etsi ratkaisut ongelmakohtiin.



Society 5.0 AI-suunnittelu

Ihminen ei toimi kuten kone.

Tulevaisuuden AI-suunnittelu perustuu tietoon siitä, miten ihminen ajattelee ja toimii.

Keskeinen ero: toinen selitysmalli.

- intentionaalinen versus kausaalinen¹.
- muitakin eroja (esim. mekaaninen vs. biologinen).

Suunnittelukysymykset voidaan ontologisoida ja organisoida, jonka avulla voidaan rakentaa työkalu².

Workflow: miten Society 5.0 AI-suunnitteluprosessi etenee

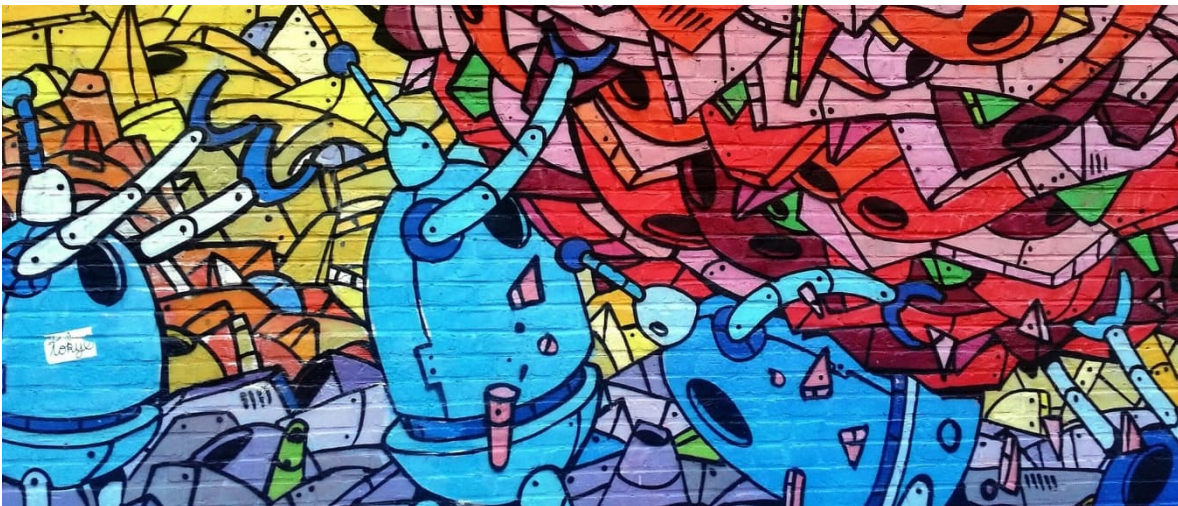
Workflow: diskurssi, mitä käydään ontologian kehittämiseksi.

Diskurssilla on omat muotonsa.

On olemassa laskennallinen prototyypityökalu (*Theseus*-työkalu, VTT).

Esim. tietyt diskurssin vaiheet:

- suunnittelukysymysten ontologian määrittely (kts. ed. sivut).
- jaetaan tehtävät, kuka ratkaisee mitkäkin kysymykset.
- kaikki kysymykset, niiden ontologia ja koko prosessi on näkyvissä
 - o esim. ”näppäimen koko kännykässä” → ratkaisuehdotus.
- ratkaisuehdotusten diskurssi (käydään läpi eri ehdotukset, vertaillaan eri ehdotuksia).
- perusteltu loppuratkaisu.
- kalenteri (ei ratkaistu, ratkaistu, to-do, WIP, jne.).



¹ Brentano, F. (1995). *Descriptive psychology*. (B. Müller, Trans. and Ed.). Routledge. (Original work published 1887)

² Saariluoma, P., Cañas, J., & Leikas, J. (2016). *Designing for life: A human perspective on technology development*. Palgrave Macmillan.

Kirjoittajat:

ETAİROS-hanke
Suunnittelun tutkimusalue

Yhteyshenkilö:

Pertti Saariluoma
ps@jyu.fi

Tutkimuskonsortio:

Tampereen yliopisto
VTT
Helsingin yliopisto
Jyväskylän yliopisto
Turun yliopisto

Rahoittaja:



ETAİROS

ETAİROS (Ethical AI for the Governance of the Society) -hanke tutkii ja etsii tapoja huomioida eettiset kysymykset tekoälyn kehittämisessä ja käytössä eri toimialoilla sekä auttaa julkista hallintoa rakentamaan tekoälyä koskevaa ohjausta ja sääntelyä.

Keskeisiä tutkimuskysymyksiä ovat:

Kuinka suunnitella eettisesti hyväksyttäviä tekoälyjärjestelmiä?
Kuinka käyttää tekoälyjärjestelmiä eettisesti hyväksyttävästi?
Kuinka säädellä tekoälyjärjestelmien käyttöä?

Tutkimushankkeen tavoitteena on vahvistaa julkisen ja yksityisen sektorin yhteistä ymmärrystä eettisten käytäntöjen ja pelisääntöjen välttämättömyydestä.

 @ETAİROSAI

