



Eettisen tekoälyn mikrokurssi

OPAS

**Osaamiskokonaisuus 5 |
Ihmisoikeudet ja
oikeudenmukaisuus**

Hankkeen numero:
2022-1-ES01-KA220-HED-000085257

Miten tätä Opasta käytetään?

Tämä dokumentti on interaktiivinen.
Dokumentissa on linkkejä lisätietoihin.



Painike, joka vie sinut dokumentin alkuun. Tämä kuvake näkyy sivujen oikeassa yläkulmassa.



Aina kun näet tämän nuolen, se tarkoittaa, että kyseessä on **interaktiivinen väriteksti**, jota voit napsauttaa ja johon on liitetty ulkoinen linkki.

VASTUUVAPAAUSLAUSEKE: Huomaa, että emme voi taata ulkoisen sisällön, kuten videoiden, jatkuvaa saatavuutta, sillä niiden tekijät tai isäntäalustat voivat muuttaa tai poistaa niitä.

Sisältö

Klikkaa valikkoa

01. Johdanto

02. Ihmisoikeuksien ja oikeudenmukaisuuden merkitys tekoälyjärjestelmissä

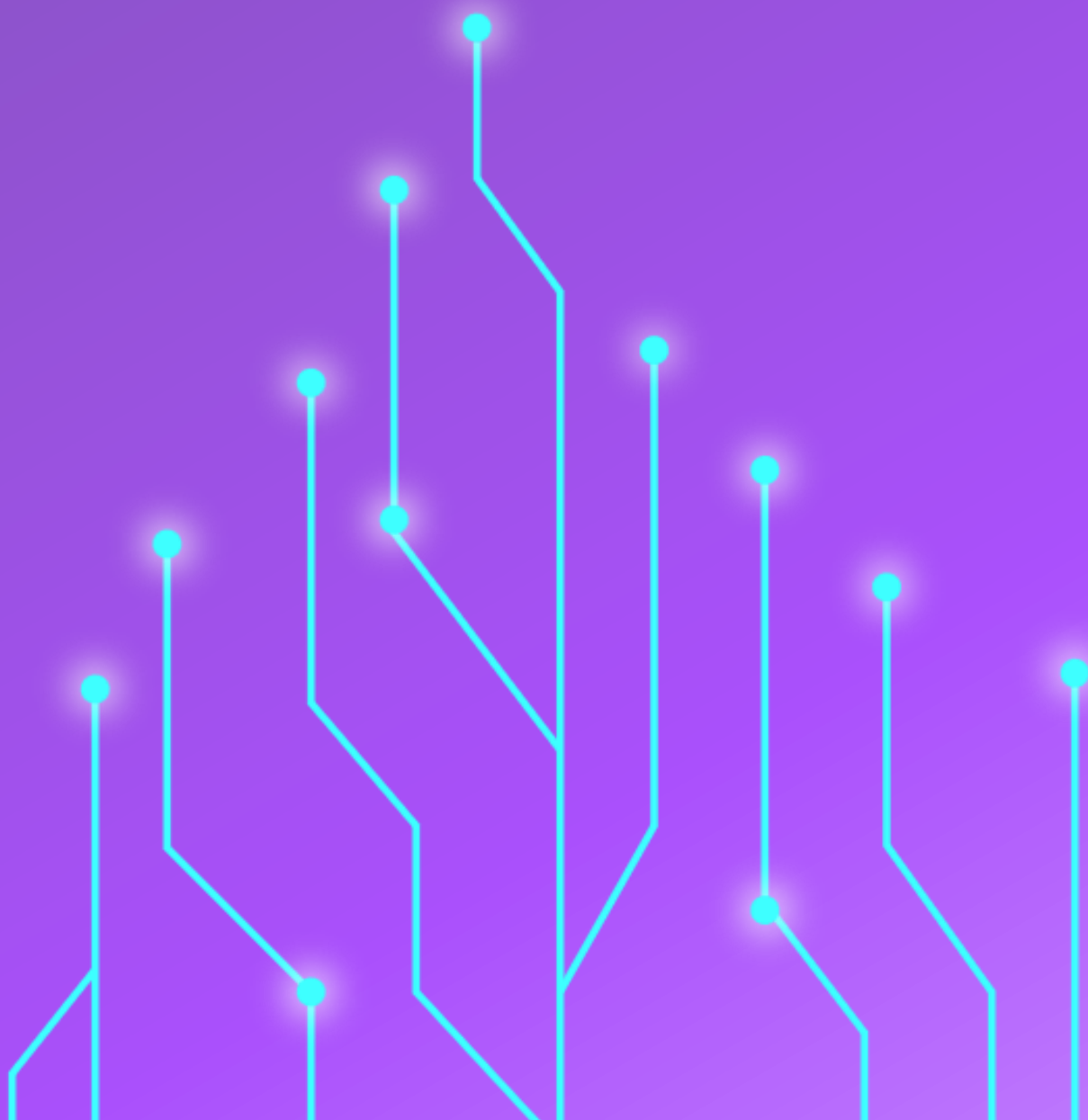
03. Ihmisoikeuksien ja algoritmisen oikeudenmukaisuuden yhtymäkohdat

04. Oikeudenmukaisuuden periaatteet tekoälyjärjestelmissä

05. Yhteenveto

01. Johdanto

Osaamiskokonaisuus 5 | Ihmisoikeudet ja oikeudenmukaisuus





01. Johdanto

Tässä osaamiskokonaisuudessa oppijat hankkivat kattavat tiedot ihmisoikeuksien ja oikeudenmukaisuuden keskeisestä roolista tekoälyjärjestelmien alalla. Tutustutaan peruskäsitteisiin, joissa korostetaan algoritmisen vinouman ja ihmisoikeuksien sekä tekoälyjärjestelmiin sisältyvien oikeudenmukaisuuden periaatteiden yhtymäkohtia. Näiden käsitteiden ymmärtämisen myötä oppijat ymmärtävät oikeudenmukaisuuden periaatteiden, tasapuolisuuden ja oikeudenmukaisuuden reaalimaailman vaikutuksia ja arvoa algoritmisen vinouman lieventämisessä ja tasapuolisempien tulosten edistämisessä ottaen asianmukaisesti huomioon tulevien sukupolvien edut.

Tämän kurssin tulokset käsittävät:

- Ihmisoikeuksien ja oikeudenmukaisuuden merkitys tekoälyjärjestelmissä oikeudenmukaisten tulosten edistämiseksi ja haittojen ehkäisemiseksi tekoälyn kehittämisessä ja käyttöönotossa. Esittelemme tekoälyteknologioihin liittyviä mahdollisia hyötyjä ja riskejä ja tunnustamme tarpeen käsitellä esimerkiksi sosioekonomisia eroja, yksityisyyden loukkauksia ja autonomian heikkenemistä.



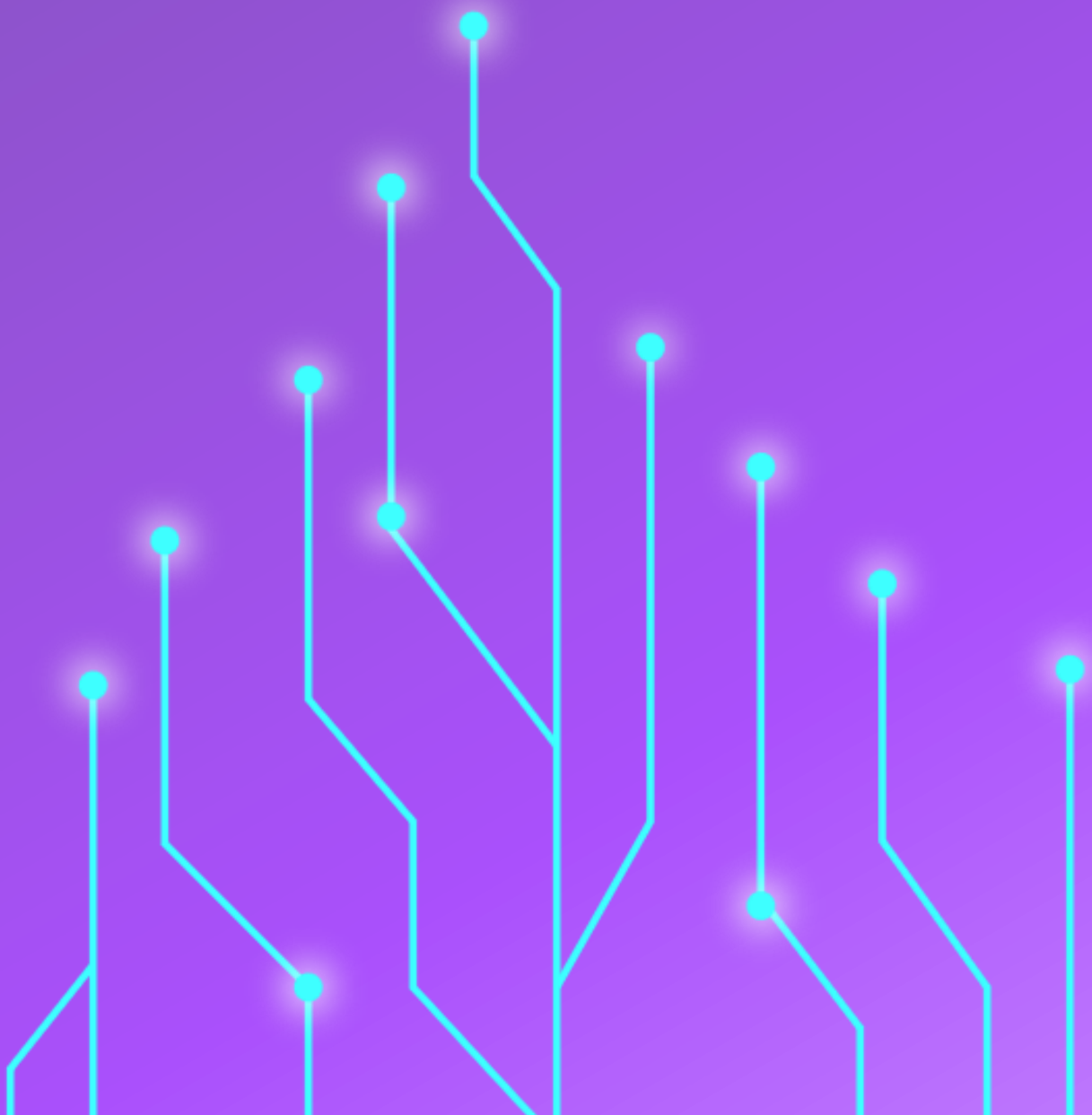


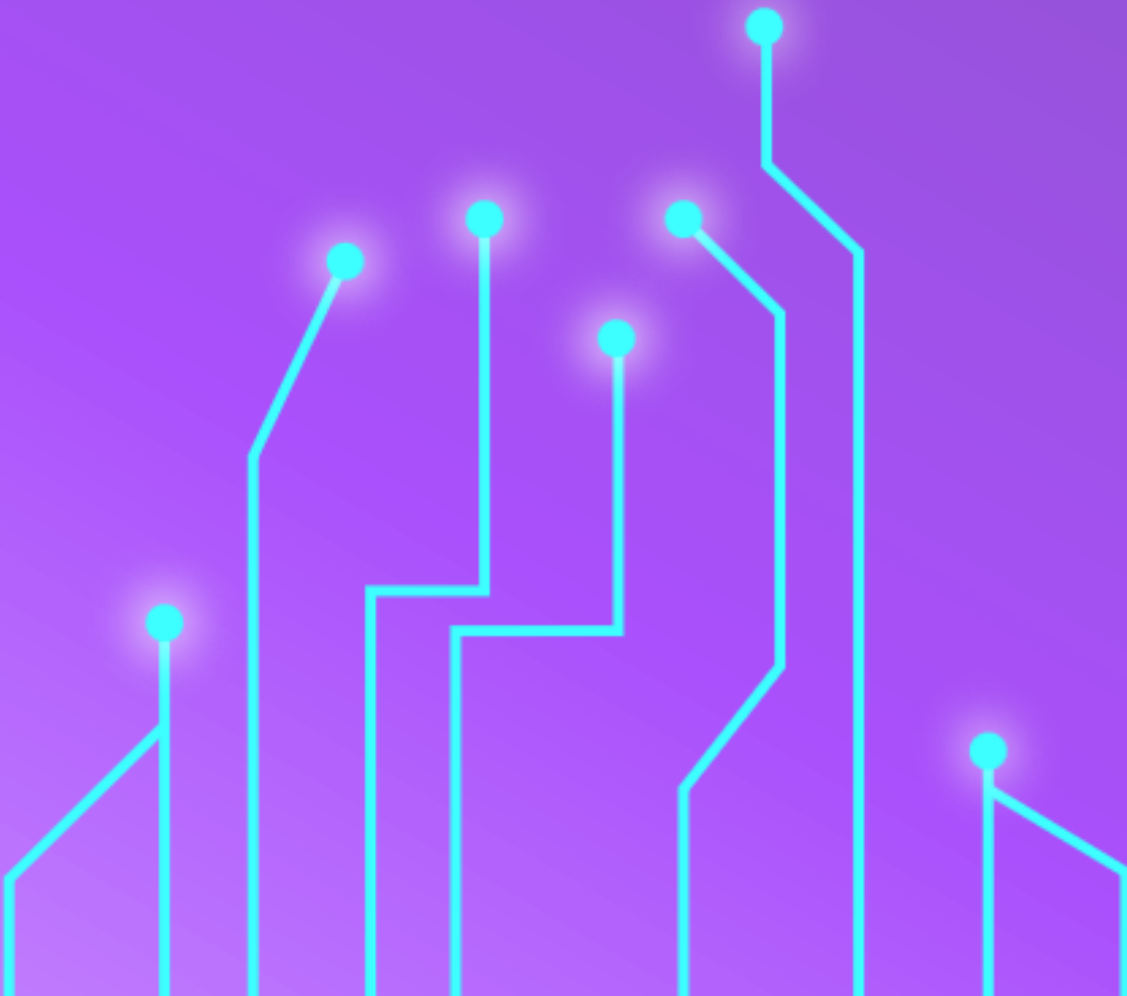
- **Ihmisoikeuksien ja algoritmisen oikeudenmukaisuuden leikkauspiste:** Tässä perusyksikössä tutustutaan ihmisoikeuksien ja algoritmisen oikeudenmukaisuuden kriittiseen risteyskohtaan ja analysoidaan teknologian eettisiä vaikutuksia teoreettisten viitekehysten pohjalta. Tutustumalla historialliseen kehitykseen ja nykypäivän keskusteluihin oppijat ymmärtävät, miten algoritmiset prosessit vaikuttavat ihmisoikeuksiin, ja tutustuvat tapaustutkimuksiin, jotka havainnollistavat algoritmisten teknologioiden ja ihmisoikeuksien välisiä ristiriitoja.
- **Ihmisoikeuksien ja algoritmisen oikeudenmukaisuuden yhtymäkohdat,** mukaan lukien yhtäläiset mahdollisuudet, syrjimättömyys, menettelyjen oikeudenmukaisuus, tasapuolisuus ja oikeudenmukaisuus. Oppijat ymmärtävät näiden periaatteiden merkityksen tulevien sukupolvien oikeudenmukaisten tulosten muovaamisessa ja saavat valmiudet selviytyä eettisistä haasteista ja puolustaa vastuullista tekoälyn kehittämistä ja käyttöönottoa.

Tekoälyjärjestelmien integrointi yhteiskunnan eri osa-alueisiin on viime vuosina herättänyt merkittäviä eettisiä huolenaiheita ihmisoikeuksien suojelusta ja oikeudenmukaisuuden edistämisestä. Tämän osaamiskokonaisuuden tavoitteena on perehtyä ihmisoikeuksien ja oikeudenmukaisuuden kriittiseen merkitykseen tekoälyjärjestelmissä ja niiden keskeiseen rooliin oikeudenmukaisten tulosten edistämisessä ja haittojen ehkäisemisessä.

02. Ihmisoikeuksien ja oikeudenmukaisuuden merkitys tekoälyjärjestelmissä

Osaamiskokonaisuus 5 | Ihmisoikeudet ja oikeudenmukaisuus





02. Ihmisoikeuksien ja oikeudenmukaisuuden merkitys tekoälyjärjestelmissä

Ihmisoikeuksien ja oikeudenmukaisuuden merkitystä tekoälyjärjestelmissä ei voi liioitella.

Ihmisoikeudet, jotka on vahvistettu useissa kansainvälisissä julistuksissa ja yleissopimuksissa, ovat eettisen hallintotavan ja yhteiskunnallisen hyvinvoinnin kulmakivi. Tekoälyjärjestelmien yhteydessä ihmisoikeusperiaatteiden soveltaminen varmistaa, että yksilöiden ihmisarvoa, itsemääräämisoikeutta ja yksityisyyttä kunnioitetaan ja suojellaan. Lisäksi tekoälyjärjestelmien oikeudenmukaisuus on olennaisen tärkeää, jotta voidaan edistää oikeudenmukaisia tuloksia ja lieventää yhteiskunnassa vallitsevien vinoumien ja eriarvoisuuden jatkumista.

Tekoälyjärjestelmät voivat vaikuttaa ihmiselämän eri osa-alueisiin, kuten työllistymismahdollisuuksiin ja palvelujen saatavuuteen, oikeudenkäyttöön ja kansalaisvapauksien suojeluun. Siksi ihmisoikeuksien turvaaminen ja oikeudenmukaisuuden edistäminen tekoälyn kehittämisessä ja käyttöönotossa ovat välttämättömiä, jotta voidaan ehkäistä haittoja ja ylläpitää eettisiä normeja. Integroimalla ihmisoikeuksia ja oikeudenmukaisuutta koskevat periaatteet tekoälyjärjestelmiin kehittäjät voivat vähentää syrjivien käytäntöjen, yksityisyyden loukkausten ja epäoikeudenmukaisten lopputulosten riskiä ja siten edistää luottamusta ja vastuullisuutta sidosryhmien välillä.



Ihmisoikeuksien ja oikeudenmukaisuuden periaatteiden soveltaminen tekoälyn kehittämisen ja käyttöönoton yhteydessä edellyttää monialaista lähestymistapaa, jossa otetaan huomioon oikeudelliset, eettiset ja yhteiskunnalliset vaikutukset. Kehittäjien on noudatettava vakiintuneita ihmisoikeuskehyksiä, kuten ihmisoikeuksien yleismaailmallista julistusta ja kansalaisoikeuksia ja poliittisia oikeuksia koskevaa kansainvälistä yleissopimusta, varmistaakseen, että tekoälyjärjestelmät noudattavat perusoikeuksia ja -vapauksia.

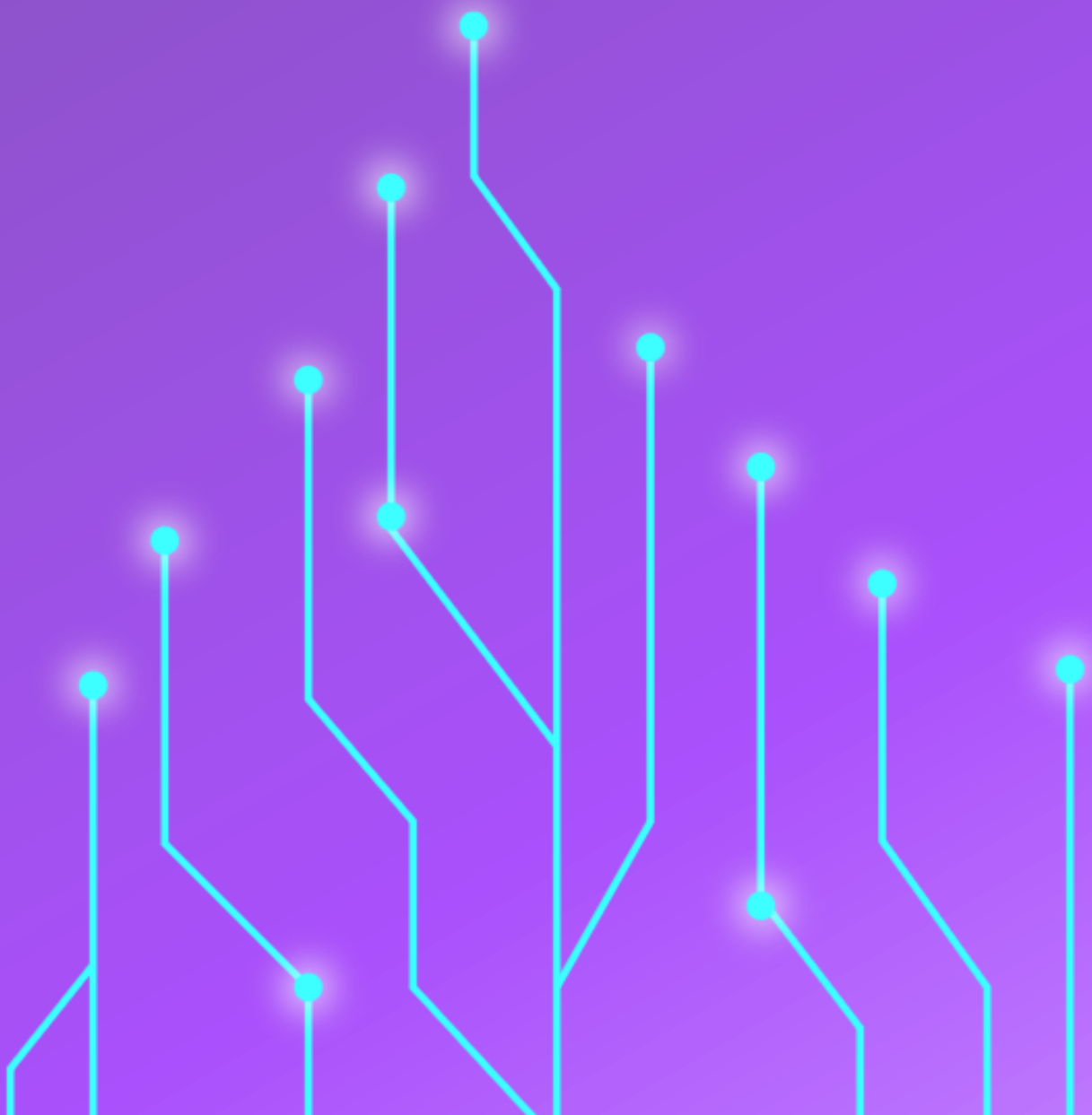
Yksi keskeinen näkökohta ihmisoikeuksien ja oikeudenmukaisuuden periaatteiden soveltamisessa tekoälyn kehittämiseen on avoimuuden ja vastuullisuuden varmistaminen koko tekoälyn elinkaaren ajan. Avoimuuden ansiosta sidosryhmät voivat ymmärtää, miten tekoälyjärjestelmät toimivat, mitä tietoja ne hyödyntävät ja millaisia päätöksentekoprosesseja niihin liittyy. Tarjoamalla selkeää dokumentaatiota ja selityksiä kehittäjät voivat antaa käyttäjille mahdollisuuden arvioida tekoälyjärjestelmien oikeudenmukaisuutta ja luotettavuutta ja saattaa niiden kehittämisestä ja käyttöönotosta vastaavat tahot vastuuseen.

Oikeudenmukaisuus tekoälyn kehittämisessä edellyttää lisäksi vinoumien lieventämisstrategioiden ja algoritmien vastuullisuusmekanismien huolellista harkintaa. Kehittäjien on ennakoivasti tunnistettava ja käsiteltävä harjoitteluaineistoon, algoritmien suunnitteluun ja päätöksentekoprosesseihin liittyviä vinoumia syrjivien lopputulosten estämiseksi. Lisäksi valvonta- ja oikeussuojamekanismien käyttöönotolla varmistetaan, että tekoälyjärjestelmien vaikutuspiiriin kuuluvat henkilöt voivat turvautua oikeussuojakeinoin, jos heitä kohdellaan epäoikeudenmukaisesti tai heille aiheutuu vahinkoa.



03. Ihmisoikeuksien ja algoritmisen oikeudenmukaisuuden yhtymäkohdat

Osaamiskokonaisuus 5 | Ihmisoikeudet ja oikeudenmukaisuus





03. Ihmisoikeuksien ja algoritmisen oikeudenmukaisuuden yhtymäkohdat

Algoritmisen vinouman ja ihmisoikeuksien välinen suhde on monitahoinen kysymys.

Algoritmisen vinouman ja ihmisoikeuksien välinen suhde on monimutkainen ja monitahoinen kysymys, joka vaatii huolellista tarkastelua. Tässä kokonaisuudessa opiskelijat syventyvät tämän suhteen yksityiskohtiin ja tunnistavat sen mahdollisen vaikutuksen haavoittuviin väestöryhmiin ja epätasa-arvon jatkumiseen. Ymmärtämällä tämän suhteen oppijat voivat tunnistaa todellisia esimerkkejä ihmisoikeuksiin vaikuttavista tekoälyjärjestelmistä ja laatia strategioita, joilla näihin kysymyksiin voidaan puuttua tehokkaasti.

Algoritmisella vinoumalla tarkoitetaan järjestelmällisiä ja epäoikeudenmukaisia mieltymyksiä tai vinoumia, joita voi esiintyä tekoälyjärjestelmien tiedoissa, algoritmeissa tai päätöksentekoprosesseissa. Jos algoritmisen vinouma jätetään valvomatta, sillä voi olla syvällisiä vaikutuksia ihmisoikeuksiin, erityisesti haavoittuvassa asemassa oleviin väestöryhmiin, kuten rotuvähemmistöihin, naisiin, vanhuksiin ja vammaisiin.

Jatkamalla olemassa olevaa eriarvoisuutta ja vahvistamalla syrjiviä käytäntöjä **vinoutuneet tekoälyjärjestelmät voivat heikentää perusihmisoikeusperiaatteita, kuten tasa-arvoa, syrjimättömyyttä ja oikeutta yksityisyyteen.**



Algoritmisen vinouman vaikutus ihmisoikeuksiin voi ilmetä eri tavoin eri aloilla ja eri yhteyksissä. Esimerkiksi rekrytointiprosesseissa käytetyt vinoutuneet algoritmit voivat johtaa syrjiviin käytäntöihin, jotka estävät yksilöitä saamasta yhtäläisiä mahdollisuuksia työhön epäolennaisten tekijöiden, kuten rodun, sukupuolen tai iän perusteella. Vastaavasti rikosoikeusjärjestelmässä vinoutuneet ennakoivan poliisitoiminnan algoritmit voivat kohdistua suhteettomasti syrjäytyneisiin yhteisöihin, mikä voi johtaa virheellisiin pidätyksiin ja loukata oikeutta oikeudenmukaiseen oikeudenkäyntiin.

Lisäksi algoritmiset vinoumat voivat pahentaa olemassa olevia eroja keskeisten palvelujen, kuten terveydenhuollon, asumisen ja koulutuksen, saatavuudessa. Luottopisteytyksessä tai lainojen hyväksymisprosesseissa käytetyt tekoälyjärjestelmät voivat systemaattisesti asettaa tietyt väestöryhmät epäedulliseen asemaan, mikä vahvistaa taloudellista eriarvoisuutta ja vaikeuttaa yksilöiden mahdollisuuksia saada taloudellisia resursseja ja mahdollisuuksia sosioekonomiseen nousuun.

Algoritmisen vinouman mahdollisen vaikutuksen tunnistaminen ihmisoikeuksiin on olennaisen tärkeää kaikkien yksilöiden, erityisesti syrjäytyneisiin tai haavoittuviin yhteisöihin kuuluvien, oikeuksien ja ihmisarvon turvaamiseksi. Analysoimalla algoritmisen vinouman ja ihmisoikeuksien välistä suhdetta oppijat saavat syvemmän ymmärryksen tekoälyteknologioiden eettisistä vaikutuksista ja tarpeesta ennakoiviin toimenpiteisiin vinoumien käsittelemiseksi ja oikeudenmukaisuuden ja tasapuolisuuden edistämiseksi.

➤ **Todellisia esimerkkejä vinoutuneista tekoälyjärjestelmistä, jotka vaikuttavat ihmisoikeuksiin, ja strategioita näiden ongelmien ratkaisemiseksi.**

Todellisia esimerkkejä ihmisoikeuksiin vaikuttavista tekoälyjärjestelmistä on paljon, mikä kertoo, että näihin kysymyksiin on puututtava kiireellisesti. **Syrjivistä kasvojentunnistusjärjestelmistä rikostuomioihin käytettäviin ennakoiviin algoritmeihin, vinoutuneet tekoälyjärjestelmät voivat loukata yksilöiden oikeuksia ja ylläpitää systeemistä eriarvoisuutta.**

Esimerkiksi lainvalvonnan yhteydessä on havaittu, että ennakoivan poliisitoiminnan ennakoivien algoritmien on todettu kohdistuvan suhteettomasti vähemmistöyhteisöihin, mikä johtaa lisääntyneeseen valvontaan, aiheettomiin pidätyksiin ja yksilöiden yksityisyyden suojan ja mielivaltaisen pidätyksen vastaisen oikeuden loukkaamiseen.

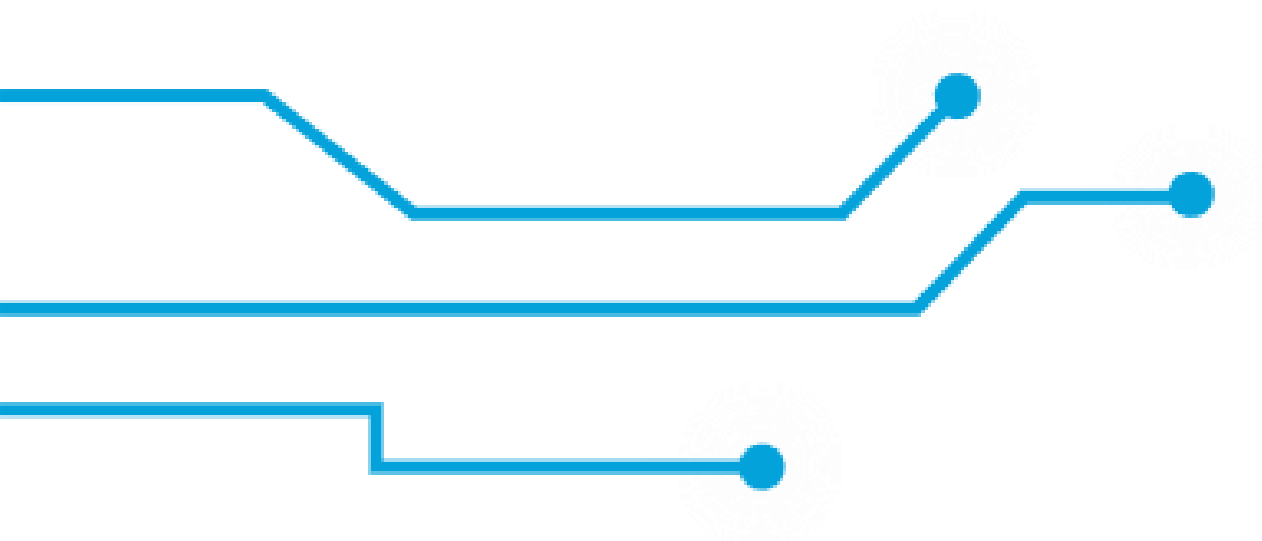
Vastaavasti terveydenhuollossa lääketieteellisessä diagnosoinnissa ja hoidon suunnittelussa käytettävät vinoutuneet algoritmit voivat johtaa väärin diagnooseihin tai riittämättömään hoitoon tietyille väestöryhmille, mikä pahentaa terveyseroja ja heikentää yksilöiden oikeutta terveyteen ja hyvinvointiin.





Näiden ongelmien ratkaisemiseksi oppijat voivat laatia strategioita, joilla pyritään lieventämään algoritmista vinoumaa ja edistämään tekoälyjärjestelmien oikeudenmukaisuutta ja tasapuolisuutta. Tähän voi kuulua vinoumien havaitsemis- ja lieventämistekniikoiden toteuttaminen kehitysvaiheessa, monipuolisen edustuksen varmistaminen koulutusdatassa sekä avoimuuden ja vastuullisuuden edistäminen algoritmisissa päätöksentekoprosesseissa.

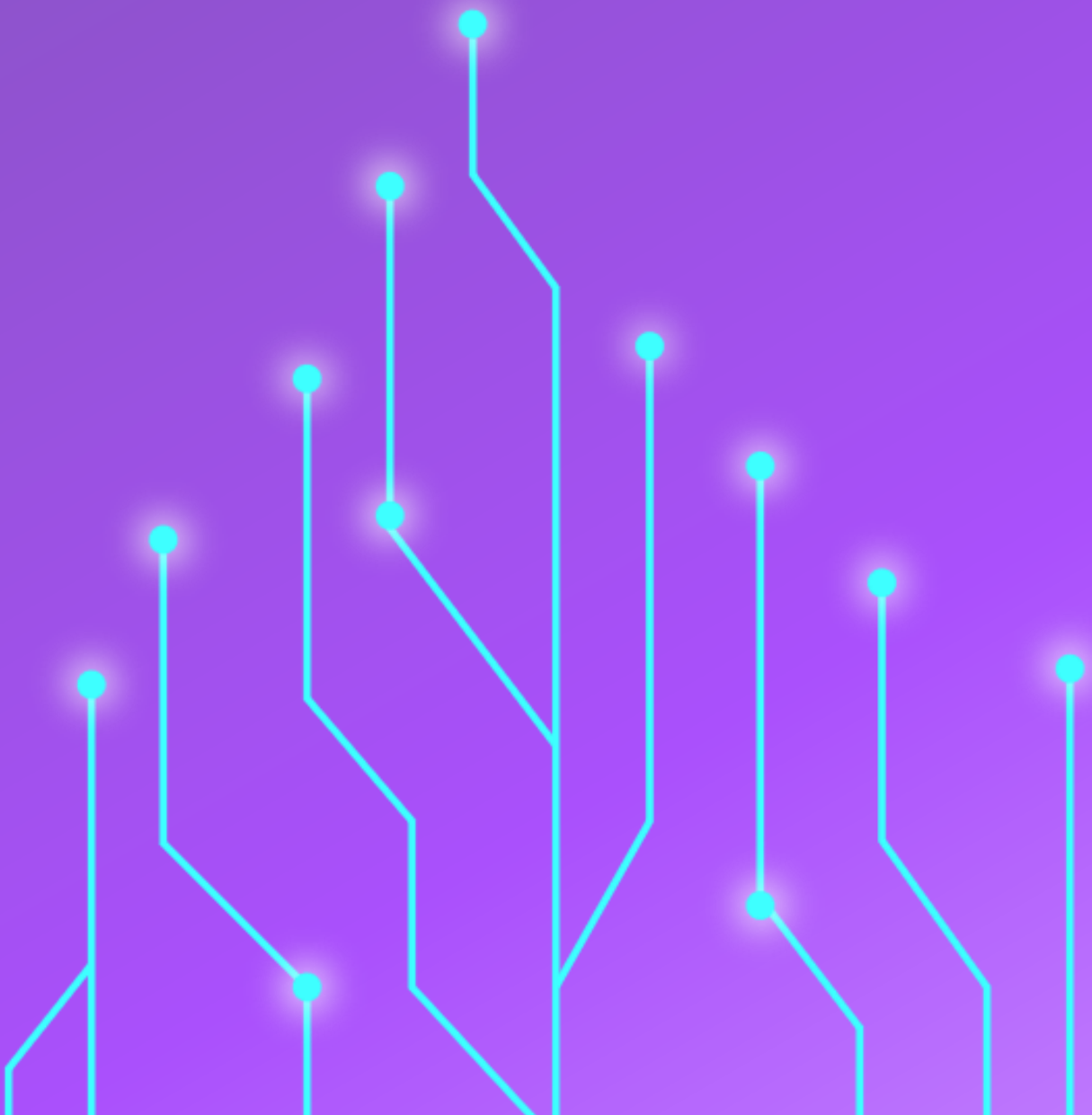
Lisäksi sidosryhmät voivat edistää sellaisten sääntelykehysten ja ohjeiden kehittämistä ja täytäntöönpanoa, joissa asetetaan etusijalle oikeudenmukaisuus, syrjimättömyys ja ihmisoikeuksien kunnioittaminen tekoälyn kehittämisessä ja käyttöönotossa.



04.

Oikeudenmukaisuuden periaatteet tekoälyjärjestelmissä

Osaamiskokonaisuus 5 | Ihmisoikeudet ja oikeudenmukaisuus





04. Oikeudenmukaisuuden periaatteet tekoälyjärjestelmissä

Oikeudenmukaisuus tekoälyjärjestelmissä käsittää useita periaatteita.

Oikeudenmukaisuus tekoälyjärjestelmissä käsittää useita periaatteita, jotka ovat olennaisen tärkeitä kaikille yksilöille tasapuolisen kohtelun ja tulosten varmistamiseksi:

- Yksi keskeisistä periaatteista on **yhtäläiset mahdollisuudet**, mikä tarkoittaa, että yksilöille tarjotaan samat mahdollisuudet heidän taustastaan tai ominaisuuksistaan riippumatta. Tällä periaatteella pyritään tasoittamaan toimintaedellytyksiä ja estämään esimerkiksi rotuun, sukupuoleen tai sosioekonomiseen asemaan perustuva syrjintä.
- **Syrjimättömyys** on toinen keskeinen periaate, joka kieltää mielivaltaisten tai epäolennaisien kriteerien käytön päätöksentekoprosesseissa. Tekoälyjärjestelmät eivät saa syrjiä yksilöitä tai ryhmiä suojattujen ominaisuuksien, kuten rodun, sukupuolen, uskonnon tai vammaisuuden perusteella. Sen sijaan niiden olisi kohdeltava kaikkia yksilöitä oikeudenmukaisesti ja puolueettomasti heidän henkilökohtaisista ominaisuuksistaan riippumatta.



- **Proseduraalisella oikeudenmukaisuudella** tarkoitetaan tekoälyjärjestelmien päätöksentekoprosessien oikeudenmukaisuutta. Siihen kuuluvat avoimuus, vastuullisuus ja oikeus valittaa päätöksistä tai kyseenalaistaa ne. Menettelyjen oikeudenmukaisuudella varmistetaan, että yksilöt saavat äänensä kuuluviin päätöksentekoprosesseissa ja että päätökset tehdään avoimesti ja vastuullisesti.
- **Oikeudenmukaisuus** on periaate, jossa keskitytään oikeudenmukaisuuden saavuttamiseen puuttumalla järjestelmään liittyvään epätasa-arvoon ja tarjoamalla resursseja ja mahdollisuuksia niitä eniten tarvitseville. Tekoälyjärjestelmien yhteydessä oikeudenmukaisuus tarkoittaa sellaisten algoritmien ja toimintatapojen suunnittelua, joissa asetetaan etusijalle syrjäytyneiden tai epäedullisessa asemassa olevien ryhmien tarpeet ja pyritään vähentämään eroja mahdollisuuksien ja resurssien saatavuudessa.
- **Oikeudenmukaisuus** on laajempi käsite, joka kattaa oikeudenmukaisuuden, tasapuolisuuden ja ihmisoikeuksien suojelun. Sillä pyritään varmistamaan, että yksilöitä kohdellaan oikeudenmukaisesti ja että heidän oikeuksiaan ja ihmisarvoaan kunnioitetaan. Oikeudenmukaisuus tekoälyjärjestelmissä edellyttää sellaisten eettisten normien, oikeudellisten säännösten ja yhteiskunnallisten normien noudattamista, jotka edistävät tasa-arvoa, oikeudenmukaisuutta ja ihmisoikeuksien kunnioittamista.

Näiden periaatteiden erottaminen toisistaan on olennaista, jotta voidaan ymmärtää tekoälyn kehittämiseen ja käyttöönottoon liittyvät monivivahteiset eettiset näkökohdat. Yhtäläiset mahdollisuudet keskittyvät yhtäläisten mahdollisuuksien varmistamiseen kaikille yksilöille, kun taas syrjimättömyys kieltää epäoikeudenmukaisen kohtelun henkilökohtaisten ominaisuuksien perusteella.

Menettelyjen oikeudenmukaisuus korostaa päätöksentekoprosessien avoimuutta ja vastuuvollisuutta, kun taas tasapuolisuus pyrkii puuttumaan järjestelmälliseen epätasa-arvoon ja edistämään syrjäytyneiden ryhmien oikeudenmukaisuutta. Oikeudenmukaisuus kattaa myös laajemmat eettiset ja oikeudelliset näkökohdat, jotka liittyvät ihmisoikeuksiin ja yhteiskunnalliseen hyvinvointiin.

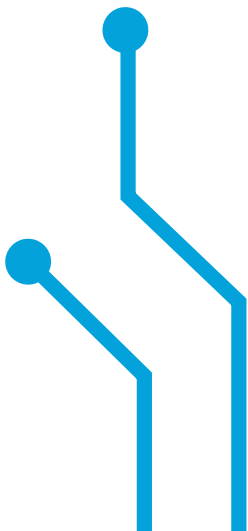
Oikeudenmukaisuuden periaatteiden soveltaminen tekoälyn kehittämiseen edellyttää ennakoivaa ja moniulotteista lähestymistapaa, jossa otetaan huomioon tekoälyteknologian eettiset, sosiaaliset ja oikeudelliset vaikutukset. Nämä periaatteet voidaan integroida tekoälyjärjestelmiin oikeudenmukaisten tulosten edistämiseksi ja algoritmisen vinouman lieventämiseksi, mikä luo pohjan oikeudenmukaisemmalle ja tasapuolisemmalle tulevaisuudelle tuleville sukupolville.



Yksi keskeinen näkökohta oikeudenmukaisuusperiaatteiden soveltamisessa tekoälyn kehittämiseen on sen varmistaminen, että algoritmit suunnitellaan ja koulutetaan käyttäen monipuolisia ja edustavia datakokonaisuuksia. Ottamalla kehitysprosessiin mukaan erilaisia näkökulmia ja kokemuksia kehittäjät voivat pienentää vinoutuneiden tulosten riskiä ja varmistaa, että tekoälyjärjestelmät ovat oikeudenmukaisia ja osallistavia.

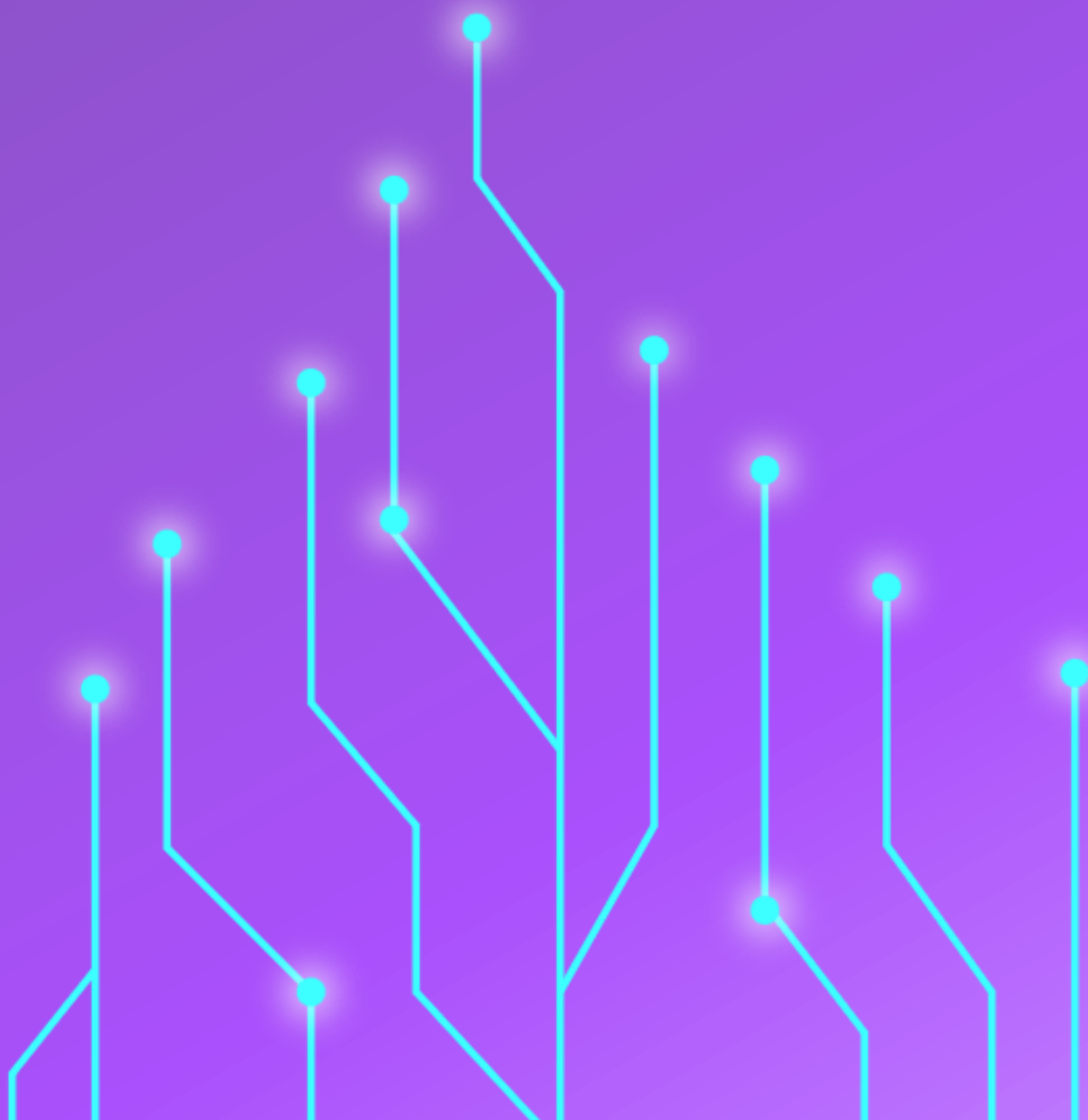
Läpinäkyvien tekoälyjärjestelmien avulla sidosryhmät voivat ymmärtää, miten päätökset tehdään, ja pitää kehittäjät vastuullisina toimistaan. Lisäksi oikeussuojakeinojen tarjoamisella varmistetaan, että henkilöillä, joihin tekoälyjärjestelmien vinouma vaikuttaa, on mahdollisuus hakea oikeutta ja korjaavia toimia.

Lisäksi tekoälyjärjestelmien oikeudenmukaisuuden edistämiseen tähtäävissä strategioissa, kuten algoritmien suunnittelussa, asetetaan etusijalle oikeudenmukaisuus ja osallisuus ja pyritään vähentämään mahdollisuuksien ja resurssien saannissa esiintyviä eroja. Ottamalla huomioon syrjäytyneiden tai heikommassa asemassa olevien ryhmien tarpeet ja näkökulmat kehittäjät voivat varmistaa, että tekoälyteknologiat palvelevat kaikkien yksilöiden etuja ja edistävät oikeudenmukaisempaa yhteiskuntaa.



05. Yhteenveto

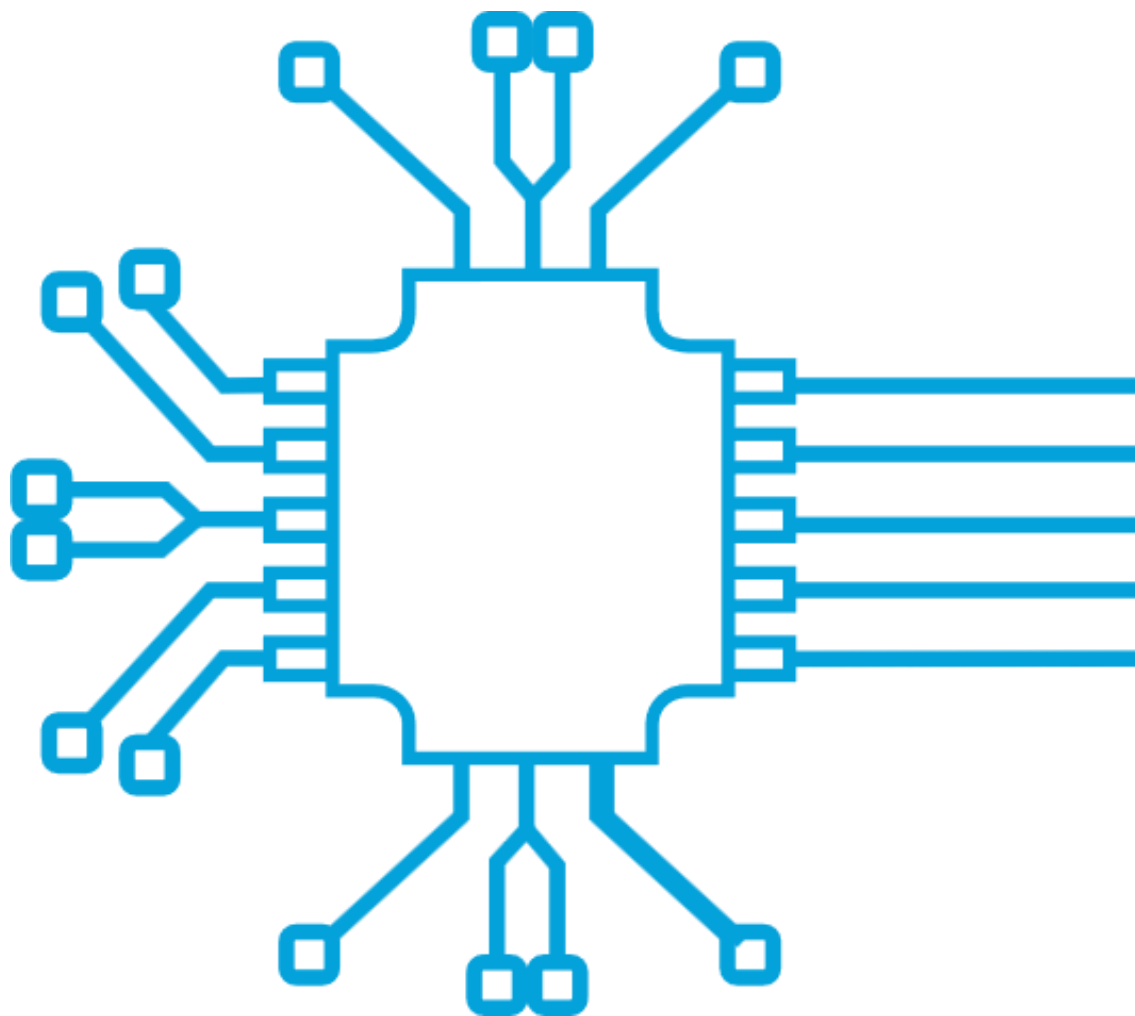
Osaamiskokonaisuus 5 | Ihmisoikeudet ja oikeudenmukaisuus

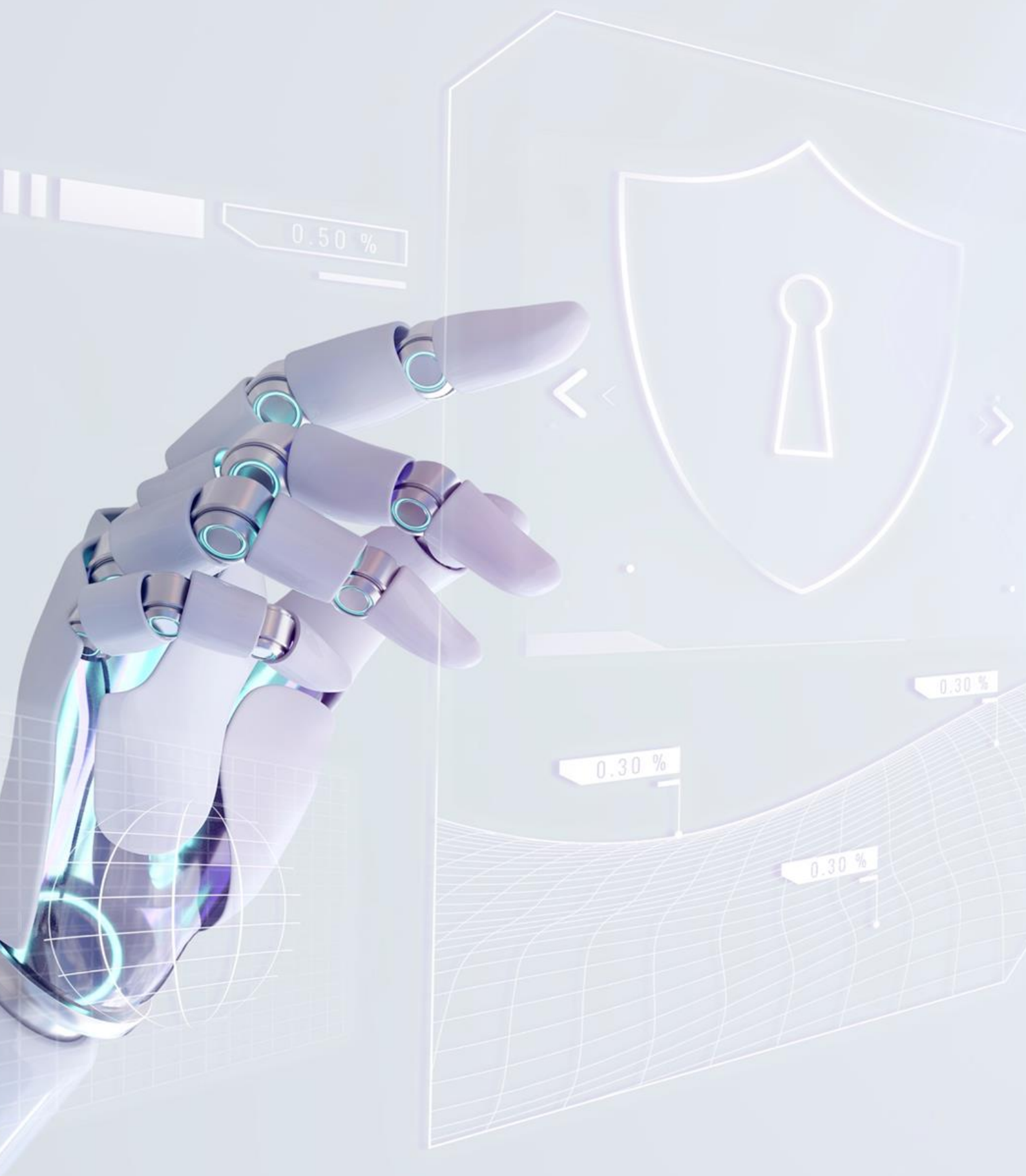




05. Yhteenveto

Yhteenvetona voidaan todeta, että ihmisoikeuksien ja oikeudenmukaisuuden periaatteiden ymmärtäminen ja soveltaminen tekoälyjärjestelmiin on ensiarvoisen tärkeää, jotta voidaan edistää oikeudenmukaisia tuloksia ja turvata perusoikeudet digitaalisena aikana.







Charlie



**Euroopan unionin
osarahoittama**

Euroopan unionin rahoittama. Esitetyt näkemykset ja mielipiteet ovat ainoastaan tämän tekstin laatijoiden näkemyksiä eivätkä välttämättä vastaa Euroopan unionin tai Euroopan koulutuksen ja kulttuurin toimeenpanovirasto (EACEA) kantaa. Euroopan unioni ja EACEA eivät ole vastuussa niistä.



**Universitat
de les Illes Balears**



helixconnect



2022-1-ES01-KA220-HED-000085257