

Tekoäly – etiikka ja tekoäly

Arvioitu oppijan taitotaso: aloittelija, keskitaso tai asiantuntija

Etiikka

Tekoälyn käyttöön liittyy eettisiä näkökulmia, kuten

- **Vinouma (bias)** (Britannica Money 2024): jotkin kategoriat voivat olla yliedustettuja tai aliedustettuja tekoälyn tuottamissa tuloksissa, erityisesti kontekstin mukaan. Kehittäjät voivat lieventää ongelmaa käyttämällä tasapainoista dataa jokaiselle kategorialle (ja kontekstille) mallien koulutuksessa.
- **Virheet tuloksissa**, esimerkiksi generatiivisen tekoälyn hallusinaatioiden tai ei-generatiivisen tekoälyn heikon suoriutumisen vuoksi sille ennestään tuntemattomalla datalla.
- **Arkaluonteinen (esim. henkilötietoihin liittyvä) tieto** – verkossa toimivat tekoälypalvelut voivat käyttää kaikkea syötettyä dataa mallien koulutukseen. Myös offline-mallit voivat “muistaa” ei-toivottuja asioita arkaluonteisesta tai luottamuksellisesta tiedosta, kuten henkilötiedoista tai yrityssalaisuuksista.
- **Tekoälyn mainitsemat lähteet ja viitteet** voivat olla olemattomia tai niistä voi puuttua olennaista tietoa. On myös mahdollista, että viitattu lähde esittää päinvastaisen väitteen kuin mitä tekoäly antaa lähteeseen perustuvana vastauksena. Lisäksi mallit voivat olla koulutettu datalla ilman asianmukaista lähdeviittausta, mikä voi johtaa plagiointiin.
- **Vastuu** – kuka on vastuussa tekoälyn tekemistä virheistä? (Britannica Money 2024) Vastuu on keskeinen näkökulma, ja tekoälyn tuottama sisältö on suositeltavaa tarkistaa ennen sen soveltamista, käyttöä tai eteenpäin jakamista.
- **Tekoälyn vastausten/tulosteiden omistajuus ja käyttöehdot** – kuka omistaa tekoälyn tuottamat tekstit, kuvat ja muut sisällöt? Saako käyttäjä hyödyntää tuotoksia vapaasti vai liittyykö niihin rajoituksia?
- **Syötedata**: onko data tuotettu eettisesti?

Tekoälysovelluksia kehitettäessä ja käytettäessä on tärkeää muistaa ja soveltaa näitä eettisiä näkökohtia.

Harjoitus: Yhdistä eettiset näkökulmat.	
Syötedata.	Luotto, lähdeviittausta, plagiointi.
Omistajuus.	Tasapaino, monimuotoisuus, yli-/ali-/vääristedustus.
Vastuu.	Anonymisointi, käytetään vain tarpeellista dataa.
Lähteet/viitteet.	Eettiset standardit koulutus- ja testidatalle.



**Euroopan unionin
osarahoittama**



**Kokkola
Karleby**

centria
University of Applied Sciences



Harjoitus: Yhdistä eettiset näkökulmat.	
Arkaluonteinen data.	Tulosten kriittinen tulkinta.
Virheet.	Tekoälystä/tuloksista vastuussa olevat henkilöt ovat tiedossa.
Vinouma	immateriaalioikeudet, oikeudet tekoälyyn/tuloksiin, lisensointi
	4,7,5,1,6,3,2

Valmis!

Onnittelten! Olet lukenut materiaalit!

Viittaukset

Britannica Money. 2024. *5 ethical questions about artificial intelligence*. Available at: <https://www.britannica.com/money/ai-ethical-issues>. Accessed 28 May 2024.



**Euroopan unionin
osarahoittama**



**Kokkola
Karleby**

CENTRIA
University of Applied Sciences

