



Līdzfinansē
Eiropas Savienība



Rīku lapa par ētisku mākslīgā intelekta dizaina kanvu

<https://www.videogames4good.eu>

Šis projekts ir finansēts ar Eiropas Komisijas atbalstu Erasmus+ programmas ietvaros. Šī publikācija atspoguļo tikai autora viedokli, un Komisija neuzņemas atbildību par jebkādu šajā publikācijā ietvertās informācijas izmantošanu. Projekta numurs: 2023-2-CZ1-KA220-YOU-000183158



YuzuPulse



LogoPsyCom



materahub

**VISAS
IESPĒJAS**



Eppas

Autors: Alan Turing Institute

Kad to pielietot:

Šis rīks ir ideāli piemērots videospēļu projekta ideju ģenerēšanas vai plānošanas fāzei, īpaši, izstrādājot spēļu mehāniku, tēlus vai sižetus, kurus varētu ietekmēt mākslīgais intelekts vai algoritmiskā loģika (piemēram, NPC uzvedība, procedūru ģenerēšana, adaptīvā grūtības pakāpe).

Tas ļaus tev attīstīt prasmes:

- Ētiskā izpratnē – ņemot vērā taisnīgumu, iekļaušanu un ietekmi spēļu sistēmās
- Sistēmiskā domāšanā – kartējot, kā spēļu mehānika un spēlētāju mijiedarbība ir savstarpēji saistīta
- Lietotājcentrētā dizainā – plānošana, ņemot vērā spēlētāju dažādību
- Sociālā atbildībā – pārdomas par to, kā spēle var ietekmēt spēlētājus un sabiedrību

Izmanto to, definējot savu mākslīgā intelekta stratēģiju spēļu veidošanā, piemēram: kā pielāgojas ienaidnieka uzvedība, kā tiek attēloti tēli vai kā dati varētu ietekmēt spēles gaitu.

Ko Tu darīsi:

Tu aizpildīsi strukturētu veidlapu, kas palīdzēs tev izvērtēt tava mākslīgā intelekta idejas ētiskās, sociālās un tehniskās sekas. Tu apsvērsi ko tas ietekmēs, kas varētu noiet greizi un kā padarīt savu dizainu taisnīgāku un drošāku.

Kas Tev būs nepieciešams:

- Drukāta vai digitāla ētiskā mākslīgā intelekta dizaina veidne (skatīt 1. pielikumu)
- Pildspalva, zīmulis vai digitālais redaktors (piemēram, Miro, Jamboard)
- Pēc izvēles: piekļuve pētniecības rīkiem (internets, raksti, piemēri)



Kā to izmantot:

1. Sāc centrā: ieraksti savu mākslīgā intelekta ideju centrālajā lodziņā (piemēram, “Mājasdarbu palīga čātbots”, “Sejas skeneris skolas drošībai”).
2. Ieinteresētās personas (augšējā labajā stūrī): Kas ir tiešie lietotāji? Kuru tas varētu ietekmēt (pozitīvi vai negatīvi)?
 - **Padoms:** Padomā par vecumu, pieejamību, pieredzi.
3. Ieguvumi un kaitējums (apakšējā labajā stūrī): Kādi ir potenciālie pozitīvie rezultāti? Kādas neparedzētas sekas varētu rasties?
 - **Padoms:** Vai tas varētu kādu izslēgt vai tikt ļaunprātīgi izmantots?
4. Aizspriedumi un taisnīgums (apakšējā centrā): vai tavš rīks varētu uzrādīt aizspriedumus (piemēram, dodot priekšroku noteiktām valodām, dzimumiem vai ādas toņiem)?
 - **Padoms:** Kā tu vari atklāt un mazināt aizspriedumus?
5. Dati un privātums (apakšējā kreisajā stūrī): kāda veida datus tas izmanto? Vai tie ir personiski vai sensitīvi?
 - **Padoms:** No kurienes tie nāk? Vai ir nepieciešama piekrišana?
6. Uzlabojumi (augšējā kreisajā stūrī): kādas izmaiņas padarītu tavu rīku ētiskāku, iekļaujošāku vai pārredzamāku?
 - **Padoms:** Vai varētu to skaidri izskaidrot lietotājam bez tehniskām zināšanām?
7. Pārskati un pārdomā: Dalies ar savu kanvu ar citiem. Saņem atsauksmes. Ko tu pārskatītu?



Par šo veidni:

Šī ir studentiem draudzīga Alana Tjūringa institūta "AI Ethics Canvas" adaptācija. Tas ir vizuālas domāšanas rīks, kas paredzēts, lai atbalstītu:

- Ideju ģenerēšanu
- Kritisku pārdomu
- Grupas diskusiju
- Iteratīvu uzlabošanu

Katra kanvas sadaļa rosina atšķirīgu domāšanas veidu::

- Sociālo (kas tiek ietekmēts?)
- Tehnisko (dati un loģika)
- Morālo (vai tas ir taisnīgi?)
- Stratēģisko (kas jāuzlabo?)

Tas pārvērš abstraktus ētikas jautājumus strukturētā aktivitātē.

Padomi atkārtotai izmantošanai/turpināšanai:

1. Pēc lietotāju testēšanas vai vienaudžu atsauksmēm atkārtoti pārskati kanvu
2. Izmanto to ar esošajiem reālās pasaules MI rīkiem, lai analizētu, cik labi tie risina ētikas jautājumus
3. Apvieno to ar prezentāciju vai īsu ziņojumu, lai pamatotu dizaina izvēles

Valodas:

Pieejams angļu, franču, čehu, latviešu un itāļu valodā.

Uzlabojumi	MI IDEJA	Ieinteresētās personas
Dati un privātums	Aizspriedumi un taisnīgums	Ieguvumi un kaitējums

