

Ethics for Human-centered AI (A005874)

Cursusomvang *(nominale waarden; effectieve waarden kunnen verschillen per opleiding)*

Studiepunten 5.0 **Studietijd 150 u**

Aanbodsessies en werkvormen in academiejaar 2026-2027

A (semester 2)	Engels	Gent	werkcollege	0.0u
			hoorcollege	0.0u
			zelfstandig werk	0.0u

Lesgevers in academiejaar 2026-2027

Daems, Joke	LW22	Verantwoordelijk lesgever
Doğruöz, A. Seza	LW22	Medelesgever

Aangeboden in onderstaande opleidingen in 2026-2027

Postgraduate Certificate Computer-Assisted Language Mediation	stptn	aanbodssessie
	5	A

Onderwijstalen

Engels

Trefwoorden

Ethiek, AI, vooroordelen en stereotypering, privacy, transparantie, verantwoordelijkheid, milieueffecten, maatschappelijke impact, eerlijkheid, mens-computerinteractie

Situering

Deze cursus belicht de risico's die gepaard gaan met het ontwikkelen en inzetten van verschillende soorten technologie en systemen die onder de (misleidende) noemer "AI" verkocht worden en onderzoekt hoe 'mensgerichte' technologie eruit zou kunnen zien. In iedere sessie gaan we aan de slag rond gevoelige thema's die door onderzoekers aangekaart worden (zoals vooringenomenheid of bias, stereotypering, gegevensprivacy en auteursrecht) en bespreken we bestaande of benodigde strategieën om deze gevoeligheden aan te pakken. De cursus benadrukt het belang van eerlijkheid, verantwoordelijkheid en transparantie (fairness, accountability, transparency) in de ontwikkeling van dergelijke systemen en is gebaseerd op recente ontwikkelingen en academisch onderzoek. De belangrijkste uitdagingen in mensgericht ontwerp zullen besproken worden. De nadruk zal vooral liggen op taaltechnologie, gezien het studieprogramma waar deze cursus deel van uitmaakt.

Inhoud

- Begrijpen hoe AI voorgesteld wordt ("framing") en wat hier de mogelijke gevolgen van zijn.
- Inzicht in de milieu- en economische impact die gepaard gaat met het ontwikkelen en onderhouden van AI-systemen. Overzicht van initiatieven om deze kosten te verlagen.
- Begrijpen wat de oorzaak is van vooringenomenheid (bias) in AI-systemen, wat de mogelijke gevolgen zijn en wat er zoals gedaan wordt om vooringenomenheid te verminderen.
- Privacy: Hoe kan de bescherming van gebruikersgegevens gewaarborgd worden en kan vertrouwelijkheid gehandhaafd worden?
- Transparantie en verantwoordelijkheid (Transparency and Accountability): Ijveren voor transparantie in taaltechnologische systemen. AI-ontwikkelaars en onderzoekers verantwoording laten afleggen voor wat ze ontwikkelen.
- Maatschappelijke impact: De invloed van taaltechnologie op de maatschappij

analyseren (bijv. misinformatie, toegankelijkheid, eerlijkheid)

Begincompetenties

De algemene competenties die van een academische bachelorstudent verwacht mogen worden.

Eindcompetenties

- 1 Uitleggen hoe de publieke perceptie beïnvloed wordt door de manier waarop er over AI gesproken wordt ("framing in AI discourses")
- 2 Mogelijke gevaren en risico's die gepaard gaan met het ontwikkelen van AI-systemen kunnen identificeren alsook de strategieën om deze risico's te beperken
- 3 De voornaamste oorzaken van vooringenomenheid (bias) en stereotypering in AI beschrijven
- 4 De impact van AI-gebruik op individuen en de maatschappij kritisch analyseren
- 5 Reflecteren over hoe taaltechnologie op een verantwoorde en ethische manier ontwikkeld en ingezet kan worden

Creditcontractvoorwaarde

Dit opleidingsonderdeel kan niet via creditcontract gevolgd worden

Examencontractvoorwaarde

Dit opleidingsonderdeel kan niet via examencontract gevolgd worden

Didactische werkvormen

Werkcollege, Hoorcollege, Zelfstandig werk

Toelichtingen bij de didactische werkvormen

De lessen zijn een combinatie van hoorcolleges voor informatieoverdracht en werkcolleges waarbinnen studenten zelf actief deelnemen. Het kan zijn dat studenten zich voor sommige lessen moeten voorbereiden door op voorhand zaken te lezen of te bekijken (wetenschappelijke artikelen en krantenartikelen, video's van lezingen en debatten, podcasts, enz., afhankelijk van relevante ontwikkelingen in het vakgebied). Om kritische reflectie te stimuleren, is deelname aan de lessen verplicht (presentatie van verwerkte materialen, bespreking van theorieën en concepten in kleinere groepen, enz.). Deze voorbereiding voor lessen en/of opdrachten is de component 'zelfstandig werk'. Er worden standaard geen lesopnames ter beschikking gesteld. Een groot deel van het vak bestaat uit groepsdiscussies, waarbij lesopnames het leerproces niet betekenisvol kunnen ondersteunen. Enkel in het geval van bijzonder statuut met faciliteit 'lesopnames' worden uitzonderingen gemaakt.

Studiemateriaal

Geen

Referenties

Referenties zullen via Ufora gedeeld worden

Vakinhoudelijke studiebegeleiding

Materiaal beschikbaar op Ufora, discussieforum voor eventuele vragen.

Evaluatiemomenten

periodegebonden en niet-periodegebonden evaluatie

Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode

Schriftelijke evaluatie

Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode

Schriftelijke evaluatie

Evaluatievormen bij niet-periodegebonden evaluatie

Vaardigheidstest, Werkstuk

Tweede examenkans in geval van niet-periodegebonden evaluatie

Niet van toepassing

Toelichtingen bij de evaluatievormen

Tijdens het semester maken studenten opdrachten om aan te tonen dat ze de vereiste competenties beheersen. De exacte opdrachten zijn afhankelijk van de specifieke onderwerpen die aan bod komen, maar dit kan bijvoorbeeld een

(Goedgekeurd)

schriftelijke samenvatting zijn van lees- en kijkmateriaal, of een presentatie in de klas over ethische kwesties die aan bod komen bij een specifieke technologie, of een lijst met 'voor'- en 'tegen'-argumenten met betrekking tot een specifieke ethische kwestie.

Er is ook een schriftelijk examen, waar studenten de kans krijgen hun begrip van de belangrijkste theoretische kaders en hun mogelijke toepassing aan te tonen.

Eindscoreberekening

Vaardigheidstest & werkstuk (opdrachten tijdens het semester): 40%

Eindexamen: 60%

De deadlines voor opdrachten moeten worden gerespecteerd. Bij een laattijdige indiening zonder voorafgaande kennisgeving aan de docent, wordt de score voor die opdracht teruggebracht tot een 9/20 (indien de student anders een voldoende zou hebben behaald).

Tijdens de tweede examenperiode moeten studenten het schriftelijk examen opnieuw afleggen. Indien de student tijdens het semester een voldoende heeft behaald voor de opdrachten tijdens het semester, blijft die deelscore behouden (deze telt dan voor 40% van het eindcijfer). Indien een student geen voldoende heeft behaald voor de opdrachten, bepaalt het schriftelijke examen 100% van de eindscore.

Faciliteiten voor werkstudenten

Aanwezigheid in de lessen wordt aanbevolen.

Beperkte mogelijkheid tot feedback via e-mail, beperkt tot het beantwoorden van specifieke vragen.