



Aan Examinatoren FGW
Van Vice-decanen FGW
C.c. Examencommissies FGW
Datum 12-06-2023
Telefoon 2094 (Anne Zwetsloot)
Nummer
Onderwerp Fraude met gebruik van chatbots (inclusief ChatGPT)

Inleiding

Deze notitie heeft als doel om uitleg te geven over de toepassing van bestaande regelgeving ten aanzien van het gebruik van chatbots (inclusief ChatGPT). Op dit moment maken veel examinatoren zich zorgen over eventuele fraude door het gebruik van chatbots, juist nu, in de tentamen- en scriptieperiode. Deze memo legt uit welke manieren er zijn om fraude door gebruik van chatbots te herkennen en welke vervolgstappen er genomen dienen te worden als dat gebeurt.

Eind november 2022 is ChatGPT gelanceerd en dit heeft de discussie over chatbots en AI in onderwijs in een stroomversnelling gebracht. ChatGPT is niet ontstaan in een vacuüm, maar parallel aan allerlei andere AI tools zoals Deepl voor vertalen en Quillbot voor parafaseren. ChatGPT trekt op dit moment de meeste aandacht omdat het gratis is en via een browser-interface voor iedereen te gebruiken is; daarnaast is er veel aandacht aan besteed in de media. Het is aannemelijk dat er op termijn meer gratis AI's beschikbaar zullen komen die vergelijkbare taken kunnen uitvoeren.

ChatGPT is een chatbot op basis van AI. Het genereert teksten als reactie op een 'prompt' (vraag of commando van een gebruiker) op basis van een enorme dataset aan tekstmateriaal (onder andere afkomstig van het internet). De gebruiker kan aangeven in welke vorm en hoe lang de tekst moet zijn. Het programma genereert deze teksten door een statistische voorspelling te doen op wat het meest waarschijnlijke volgende woord in de tekst zou zijn. Die voorspelling is gebaseerd op zinsdelen of eerder gebruikte zinnen. Chatbots begrijpen niet wat ze schrijven: de kwaliteit van het antwoord is direct gerelateerd aan de kwaliteit van de vraag (de prompt).

Chatbots zoals ChatGPT hebben allerlei consequenties voor het onderwijs, zowel positief als negatief. Meer informatie over chatbots in het onderwijs vind je op de websites genoemd aan het eind van deze memo.



1. Wanneer is het fraude?

Wanneer een student (in grote mate) gebruik maakt van AI voor een schrijfoopdracht en de beoordelaar geen onderscheid kan maken tussen de delen die door de student of de AI geschreven zijn, dan belemmert dat de oordeelsvorming en is er sprake van fraude. Onder fraude wordt in de zin van artikel 7.12b van de Wet op het Hoger Onderwijs (WHW) verstaan ieder handelen of nalaten van een student dat erop is gericht het vormen van een juist oordeel omtrent hun kennis, inzicht en vaardigheden geheel of gedeeltelijk onmogelijk te maken. Hieronder valt, onder andere, het door anderen laten maken van (delen van) een tekst.

2. Hoe herken je het gebruik van chatbots?

Het kan lastig zijn om exact vast te stellen of er gebruik is gemaakt van chatbots, maar er zijn wel een aantal aanwijzingen:

- **Schrijfstijl:** Als de schrijfstijl in een tekst (sterk) afwijkt van de gebruikelijke schrijfstijl van de student. Daarnaast kan het, volgens ChatGPT zelf, worden herkend aan de volgende elementen:
 - Herhalende of langdradige formuleringen
 - Onlogische of inconsistente antwoorden
 - Overmatig gebruik van specifieke woorden of uitdrukkingen
 - Overmatig gebruik van adjectieven en bijvoeglijke bijzinnen
 - Gebrek aan persoonlijke flair of individualiteit
 - Beperkte variatie in zinsstructuur
 - Minder gebruik van informele taal of spreektaal
 - Ongebruikelijke of onnatuurlijke woordkeuze
- **Inhoudelijke fouten:** Een chatbot is een taalmodel en kan dus niet zelf nadenken. De chatbot kan onwaarheden genereren met een sterke overtuiging. Dit komt doordat de software is getraind op enorme hoeveelheden data afkomstig van het internet, waar ook onjuiste informatie te vinden is.
- **Verouderde data:** De huidige versie van ChatGPT (januari 2023) is alleen getraind op data tot en met september 2021. Dit betekent dat de chatbot recentere data niet zelf zal gebruiken. Deze beperking zal waarschijnlijk van korte duur zijn, gezien de snelheid van de ontwikkelingen. Daarnaast kan een student wel actuele data invoeren in een prompt, die ChatGPT dan kan gebruiken. Ook kan ChatGPT wel worden gedwongen om actuele data weer te geven: hij zal deze dan verzinnen op basis van oudere teksten.
- **Rekenfouten:** Het is een taalmodel, geen rekenmachine: tekst wordt gegenereerd als antwoord op mathematische vragen, maar de uitkomst kan een willekeurig getal zijn op basis van menselijke voorkeur voor getallen.
- **Incorrecte bronvermelding:** Chatbots genereren antwoorden aan de hand van kansberekening, gebaseerd op talloze bronnen. Dit gebeurt bij het opstellen van stukken tekst, maar ook bij het opstellen van een simpel feitelijk antwoord op een vraag. Dit betekent dat je nooit precies kunt nagaan op welke bronnen



een chatbot zich baseert. Dit maakt het antwoord van een chatbot niet transparant, moeilijk te controleren en maakt een juiste en volledige bronvermelding onmogelijk.

- **Niet bestaande bronnen:** Als je ChatGPT vraagt de bronnen te geven die gebruikt zijn voor het genereren van een antwoord, dan fabricceert de chatbot vaak niet bestaande bronnen. Je kan vragen om bestaande bronnen of bronnen met peerreview, dan krijg je meestal wel toepasselijke bronnen, maar deze bronnen zullen niet alle data bevatten waarop de chatbot een antwoord baseert. Het gebruik van niet bestaande bronnen kan worden gezien in de plagiaatcontrole van Turnitin: normaliter worden bronnen geel gearceerd omdat het bestaande tekst is, maar bij gefabriceerde bronnen zal Turnitin de tekst beschouwen als origineel en dus niet markeren.

3. Turnitin: (nog) niet betrouwbaar en wordt verwijderd

De uitingen van de chatbot zijn uniek; de teksten zijn niet op internet terug te vinden. De standaard plagiaatcheckers werken dan ook niet. De Universiteit Leiden maakt op dit moment geen gebruik van software die kan opsporen of er gebruik is gemaakt van een chatbot omdat de resultaten (nog) niet betrouwbaar genoeg zijn; dit geldt ook voor de tool in Turnitin. De plagiaatcontrole van Turnitin is gewoon betrouwbaar, net als altijd. Dit percentage kan dan ook gebruikt worden.

Vanaf 16 juni jl. zal de AI Writing Detection Tool in Turnitin worden uitgezet. Sinds begin april heeft Turnitin een AI Writing Detection tool geïntegreerd in FeedbackStudio. Ondanks dat Turnitin zelf claimt dat de kans <1% is dat deze tool een tekst markeert als “geschreven door AI” terwijl dat niet het geval is, kunnen wij deze claim niet verifiëren of reproduceren. Het is ook niet te achterhalen hoe de Turnitin tool precies werkt en waar deze zijn conclusies op baseert. Daarom is er op centraal niveau besloten om deze functionaliteit voor nu uit te zetten in Turnitin en zal deze ook helemaal niet meer zichtbaar zijn; dit geldt voor bestaande en toekomstige submissions. Deze rapportages zijn ook niet meer op te vragen.

4. Wat te doen bij het vermoeden van chatbot gebruik?

- **Meld je vermoeden:** Examinatoren zijn verplicht (een vermoeden van) fraude te melden bij de examencommissie.
 - Stuur een e-mail naar de examencommissie met het bewijsmateriaal in de bijlage
 - De examencommissie heeft nodig: 1) screenshots van alle fragmenten waar je AI gebruik vermoedt; 2) de Turnitin plagiaat-rapportage; 3) de originele inzending
 - Ga niet in discussie met de student en deel de screenshots niet. Stel de student op de hoogte dat de zaak wordt voorgelegd aan de examencommissie en dat deze alle vragen zal beantwoorden. Het is mogelijk om een opmerking in Brightspace toe te voegen dat het paper nog niet is beoordeeld omdat er gewacht wordt op een uitspraak van de



examencommissie en deze opmerking zichtbaar te maken voor de student.

- Geef de examencommissie in je e-mail een leidraad: beschrijf of de opdracht voldoet aan de vereisten en waarom je denkt dat de opdracht door AI is gegenereerd.
- Verifieer voordat je de case voorlegt eerst of de referenties bestaan. Let op: dit kan ook zijn dat er verwezen wordt naar relevante, bestaande en bekende auteurs waarvan geen artikelen of studies onder een dergelijke titel bestaan. Zij hebben de publicatie dus niet geschreven. Geef in je e-mail aan of je de referenties hebt gecontroleerd en of ze echt zijn.
- **Geef (nog) geen cijfer of beoordeling:** Bij een vermoeden van fraude wordt een gemaakt tentamen (opdracht/verslag) pas door de examiner beoordeeld en van een cijfer voorzien wanneer de examencommissie een besluit heeft genomen over het al dan niet daadwerkelijk plaatsvinden van fraude en over het al dan niet opleggen van sancties.
- **Leg geen sancties op:** Wettelijk is bepaald dat sancties (zoals het ongeldig verklaren van de gemaakte opdracht of uitsluiten van het tentamen) alleen mogen worden opgelegd door de examencommissie. Dat mag dus niet door examinatoren.

5. Wat doet de examencommissie?

De examencommissie zal zaken betreffende fraude door het gebruik van AI tekst-generatoren hetzelfde behandelen als andere zaken omtrent fraude en plagiaat. De examencommissie zal bepalen of er sprake is van fraude door het gebruik van AI, al dan niet in gesprek met de student. De examencommissie dient in haar besluit, naast de opgelegde sanctie, ook duidelijk te vermelden waaruit de overtreding bestond en wat het overtreden voorschrift was. Daarbij moet ook duidelijk verwezen worden naar bijvoorbeeld de Regels en Richtlijnen of OER. In het geval dat de examencommissie fraude door een student vaststelt, legt zij sancties op. Deze sancties staan in verhouding tot de aard en zwaarte van de gepleegde fraude.

6. Afsluiting

Het gebruik van AI in onderwijs blijft een lastig onderwerp en het is belangrijk om het onderwerp van wetenschappelijke integriteit te bespreken met studenten, ook specifiek in relatie tot AI. De universiteit verbiedt het gebruik van chatbots niet, maar er is o.a. in de handleiding voor examencommissies opgenomen dat de tekst van de student niet door een ander mag zijn geschreven, en dat is inclusief een chatbot. Doen alsof een AI-gegenereerde tekst een eigen, originele tekst is, is in strijd met de richtlijnen voor wetenschappelijke integriteit.

We horen op dit moment veel signalen over een toename van fraudegevallen als gevolg van het gebruik van ChatGPT. We snappen dat het op dit moment een flinke belasting vormt voor examinatoren en examencommissies. Het is belangrijk dat we studenten goed



**Universiteit
Leiden**

Memo

informereren over het gebruik van AI om dit zo veel mogelijk te voorkomen, maar dat neemt de extra belasting op dit moment niet weg. We hebben daar zorgen over en daarom monitoren we de ontwikkelingen binnen de faculteit en universiteit alsmede de ontwikkelingen op gebied van AI zo goed mogelijk.



Contactinformatie voor ondersteuning en vragen:

- ECOLe (ecole@hum.leidenuniv.nl) voor vragen over Turnitin
- O&K (oenk@hum.leidenuniv.nl) voor vragen over deze memo

Meer informatie:

Wil je meer lezen over AI in het onderwijs?

- <https://www.medewerkers.universiteitleid.nl/onderwijs/ict-en-onderwijs/ai-in-het-onderwijs/ai-in-het-onderwijs>
- de FGW ChatGPT leeslijst:
<https://www.medewerkers.universiteitleid.nl/vr/geesteswetenschappen/fgw-onderwijs/tips-en-tricks/chatgpt-leestips-en-tools>
- <https://teachingsupport.universiteitleid.nl/getting-started/on-chatgpt>
- <https://www.leidenlearninginnovation.org/stories/the-possibilities-and-problems-of-ai-and-its-detection-in-education/>

Ervaring opdoen? Hier kun je oefenen met ChatGPT via LLinC:

- <https://www.leidenlearninginnovation.org/stories/chatgpt-exercise/>

Of volg de cursus over ChatGPT tijdens de onderwijsparade op 30 juni 2023:

- <https://www.medewerkers.universiteitleid.nl/agenda/2023/06/onderwijsparade-2023?cf=geesteswetenschappen&cd=fgw-bestuur-bureau>

Over Machine Translation:

- [Machine Translation Literacy \(google.com\)](#)
- [Machine Translation Literacy - Teaching Resources \(google.com\)](#)

Collega's bij FSW hebben een blog geschreven over de AI Detection tool. Lees hem hier:

- <https://www.leidenanthropologyblog.nl/articles/when-all-positives-are-false-the-emergence-of-an-academic-whodunnit>

Meer informatie over de Turnitin AI Writing Detection tool is hier te vinden:

- <https://universiteitleid.nl/screensteps/live.com/a/1666201-ai-writing-detection>.

Over AI/ChatGPT algemeen:

- <https://lifearchitect.ai/>