

Handreiking (G)AI voor examencommissies en examinatoren

De Haagse Hogeschool

let's change
YOU. US. THE WORLD.

Auteurs

mr. Corine van Toor-van Essen en drs. Joyce den Heijer

Versie april 2024

DE HAAGSE
HOGESCHOOL

Toelichting bij de handreiking (G)AI voor examencommissies

Generatieve AI is een specifieke vorm van artificial intelligence. Het helpt bij het genereren van tekst, computercode, beeld en audio en is nu al een krachtig verlengstuk van het analytisch en scheppend vermogen van mensen gebleken. Zeker sinds de lancering van ChatGPT is het gebruik van artificial intelligence onderdeel gaan uitmaken van het leven van veel mensen (aldus de Aanbiedingsbrief overheidsbrede visie op generatieve AI van 18 januari 2024).

Eind november 2022 ontstond in de vorm van ChatGPT een nieuwe uitdaging op het gebied van onderwijs en toetsing: een toegankelijke chatbot die op het eerste oog overtuigende teksten en beroepsproducten kan opleveren. Sinds 2023 is de aandacht voor Generatieve Artificial Intelligence (hierna afgekort als (G)AI) alleen maar toegenomen. Het blijkt een handige en veel gebruikte tool voor studenten, docenten en andere onderwijsprofessionals. Alle betrokkenen moeten leren omgaan met (G)AI. Dit raakt het onderwijs, de toetsing en daarmee het werk van studenten, docenten en examencommissies. Van belang is het borgen van de kwaliteit van de toetsing, maar ook het doordenken van toepassingen van AI in onderwijs en toetsing. Er zijn kansen en risico's.

Je kunt [hier](#) uitleg vinden wat de kansen en risico's concreet betekenen voor het gebruik van (G)AI en de kernwaarden van De Haagse Hogeschool, te weten:

- a. connecting: AI literacy als voorwaarde voor digitale literacy;
- b. caring, welzijnsgericht en verantwoord faciliteren van AI;
- c. curious: voortrekkersrol in toegepaste AI als voorwaarde voor relevant onderwijs.

In de [Onderwijsvisie](#) is de hogeschoolbrede afspraak opgenomen: "Iedere opleiding leert studenten verantwoord om te gaan met AI en digitale tools en zet die doelbewust in, zodat studenten daarin vaardig worden en ook op dit gebied goed voorbereid zijn op de toekomst. Iedere opleiding werkt volgens het hogeschoolbrede AI-beleid". In maart 2024 is een hogeschoolbreed (G)AI beleidskader door het College van Bestuur vastgesteld.

De examencommissie heeft de taak en bevoegdheid om (onder meer) **fraude** vast te stellen en te sanctioneren en de **kwaliteit van toetsen en examens te borgen**. Deze handreiking AI voor examencommissies gaat in op deze taken en bevoegdheden en vormt een aanvulling op de

1. handreiking voor studenten (academische integriteit);
2. handreiking (G)AI en toetsing;

beide geschreven door de collega's van het Toetslab en onderdeel van de [Toolkit AI in je toetsontwerp](#) (zie ook paragraaf 2.3, waar ook de links naar deze documenten staan).

Deze handreiking is tot stand gekomen met raadpleging van de examencommissies, het juridisch team, team Onderwijsadvies: Frank van Nispen en Janine Haenen (OKC) en team Kwaliteit en Beleid (OKC), Sascha Winkel (IT&D) en vertegenwoordiging van de [Versnellingsagenda AI](#). In verband met de snelle ontwikkelingen van (G)AI en de consequenties voor de onderwijs- en toetspraktijk zullen de ontwikkelingen nauwgezet worden gevolgd en zal deze handreiking regelmatig worden geëvalueerd en zo nodig worden aangevuld.

In deze handleiding wordt 'hij' en 'zijn' gebruikt als (de) student(en) of (de) examiner(en) wordt bedoeld.

Inhoudsopgave

Inhoudsopgave.....	3
1. Het vaststellen en sanctioneren van AI-plagiaat	4
1.1 Plagiaat met (G)AI in OER 24-25	4
1.2 AI-plagiaat voorkomen	4
1.3 AI-plagiaat herkennen	5
1.4 AI-plagiaat bewijzen: wat zegt de Raad van State?	5
1.5 AI-plagiaat bewijzen: een rol voor de examinerator	6
1.6 AI-plagiaat bewijzen: de taak van de examencommissie [stappenplan]	7
1.7 Handvatten uit een uitspraak van het College van Beroep voor de Examens (CBE)	8
2. Het borgen van de kwaliteit van toetsen en examens.....	9
2.1 Taken en verantwoordelijkheden	10
2.2 Richtlijnen van de examencommissie	10
2.3 Kwaliteitsprincipes toetsing en (G)AI	10
3. Casuïstiek uit de praktijk	14
4. Bronnen	17

1. Het vaststellen en sanctioneren van AI-plagiaat

1.1 Plagiaat met (G)AI in OER 24-25

In de Onderwijs- en Examenregeling (OER) 24-25 is plagiaat met behulp van artificial Intelligence toegevoegd aan de bestaande bepaling over plagiaat.

In artikel 7.1 lid 3 van OER 24-25 is bepaald dat plagiaat een vorm is van fraude. Door het werk van een ander te gebruiken *of door gebruik te maken van artificial intelligence (AI)* is degene die het werk beoordeelt niet in staat het kennen en kunnen van de student op een correcte en eerlijke manier te beoordelen.

De volgende situaties worden aangemerkt als plagiaat:

- het letterlijk of in eigen woorden verwerken van delen van het werk van een ander, al dan niet met aanhalingstekens, in eigen werk zonder verdere specificatie of bronvermelding.
- het presenteren van uitgewerkte ideeën of vondsten van anderen als eigen ideeën of vondsten.

Als de student (G)AI geproduceerd werk inlevert, zonder bronvermelding en het presenteert als eigen werk is dat dus plagiaat (verder: AI-plagiaat). Voorheen was het bewijzen van plagiaat relatief eenvoudig met behulp van de Original plagiatscan. De feiten en omstandigheden die tot de conclusie leiden dat plagiaat is gepleegd waren dan duidelijk. Veel lastiger is het om in geval van ongeoorloofd gebruik van AI vast te stellen dat sprake is van plagiaat: er is op dit moment geen (binnen De Haagse Hogeschool goedgekeurde) AI-detectietool beschikbaar.



Opmerking vooraf. Het is van belang dat opleidingen toewerken naar AI-resilient ('veerkrachtige') toetsprogramma. Naast deze handreiking geeft ook de toolkit AI-toetsontwerp daarvoor verschillende handvatten en concrete oplossingsrichtingen (zie paragraaf 2.3).

Met name in de periode dat dit nog niet is gerealiseerd zal de detectie en het onderbouwen van vermoedens van ongeoorloofd gebruik van (G)AI aandacht en inzet vragen van examencommissies en examinatoren. In deze handreiking wordt een aantal handvatten geboden.

1.2 AI-plagiaat voorkomen

Het is in het belang van alle betrokkenen om fraude te (helpen) voorkomen. Naast het gebruik van geschikte toetsvormen om de kans op fraude te beperken, is duidelijke communicatie met de student nodig over het (wel of niet toegestaan) gebruik van (G)AI. Het kan zijn, afhankelijk van de leeruitkomsten van de onderwijseenheid/module, dat het de student wordt toegestaan om (al dan niet in beperkte mate) gebruik te maken van AI, of helemaal niet.

Communicatie aan de student

Voor de start van de onderwijseenheid/module moet in elke modulebeschrijving beschreven zijn welk gebruik van (G)AI al dan niet is toegestaan. Het is aan te raden om daarvoor een uniforme formulering of sjabloon te gebruiken. Dit bevordert de consistentie en duidelijkheid voor studenten. Dit valt onder de zorgplicht en verantwoordelijkheid van de opleiding. Als dit nog niet in modulebeschrijving is opgenomen, is het van belang dat de docent dit niet alleen mondeling bespreekt met de studenten, bijvoorbeeld tijdens de eerste bijeenkomst, maar dit ook schriftelijk, via Brightspace, onder de aandacht brengt, zodat alle studenten goed geïnformeerd zijn.

Naast schriftelijke bevestiging is het aan te raden om de communicatie over (G)AI-richtlijnen ook vast te leggen in een centraal toegankelijk document of platform voor zowel docenten als studenten, zoals het intranet, Brightspace of een FAQ-sectie over (G)AI-gebruik.

1.3 AI-plagiaat herkennen

Het is niet eenvoudig fraude met AI te herkennen bij geschreven tekst, visualisaties en programmeercodes. Bij geschreven tekst zijn er een aantal aanwijzingen:

- Het kan nuttig zijn om de schrijfstijl in het stuk te vergelijken met eerder werk van de student. Bij discrepanties in de verschillende werken van de student kan dat een aanwijzing zijn dat de student gebruik heeft gemaakt van AI. Aan te raden is om navraag te doen bij de examinerator of studentcoach die een goed beeld heeft van eerder werk van de student.
- Een aantal aanwijzingen die kunnen leiden tot het vermoeden dat een tekst door een chatbot is geschreven (Chatten met Napoleon, B. Last (2023), blz. 85):
 - Ontbreken van correcte citaten waar deze vereist of verwacht worden;
 - Onjuist of inconsistent gebruik van eerste- en derdepersoonsperspectief;
 - Opvallen gebruik van anglicismen;
 - Taalgebruik of vocabulaire dat niet past bij het niveau van de student;
 - Geen verwijzingen naar gebeurtenissen na een bepaalde datum, die samenhangt met de dataverzameling van een specifieke AI-tool;
 - Overdreven gebruik van overdreven of formeel taalgebruik;
 - Waarschuwingen of voorbehouden die gegenereerd zijn door (G)AI, die bijvoorbeeld de beperkingen of hypothetische aard ervan benadrukken.
- De chatbot kan inhoudelijke fouten en rekenfouten genereren omdat het een taalmodel is en dus niet zelf kan nadenken. De kans op onjuiste informatie is zeker niet verwaarloosbaar.
- Het gebruik van abstracte ideeën waarvan de docent vermoedt dat de student ze niet begrijpt of kan toepassen.
- Het stuk dat door een chatbot wordt geproduceerd is vaak niet transparant als het gaat over bronvermelding. ChatGPT creëert zowel bronnen die niet bestaan als bronnen die qua inhoud niet overeenkomen met hetgeen het stuk naar verwijst. Aan te raden is dus om een check te doen op de (inhoud van de) bronnen.
- De inhoud van werkstukken van verschillende studenten komt sterk overeen, vanwege op elkaar lijkende output die door (G)AI wordt gegenereerd.

Ten overvloede: het is niet altijd zo dat als werk van de student een of meer van deze kenmerken vertoont, er sprake is van AI-plagiaat.



Op dit moment is het niet toegestaan om een AI detectietool te gebruiken om plagiaat op te sporen, vanwege de intellectuele eigendomsrechten van studenten. Bovendien is het uploaden van het werk van studenten in (G)AI zonder expliciete toestemming van de student in strijd met [Privacyverklaring](#) van De Haagse Hogeschool.

Uit inventarisatie van ervaringen van opleidingen blijkt dat er wel degelijk AI-detectietools worden ingezet als een hulpmiddel om te komen tot een onderzoek naar AI-plagiaat. Dat is dus niet toegestaan. Daarbij komt dat scores die worden gegenereerd door AI-detectietools niet geschikt zijn om te bewijzen dat een student zich schuldig heeft gemaakt aan plagiaat. De scores zijn onbetrouwbaar (hoog risico op onjuiste en vals-positieve scores en gebrek aan transparantie over hoe de tools werken, wat van invloed kan zijn op o.a. het intellectuele eigendomsrecht). Als je hier meer over wilt weten kun je deze [link](#) raadplegen. Bovendien zijn er tools om een tekst 'undetectable' te maken en zijn premium (duurdere) AI tools slechter detecteerbaar (wat ongelijkheid met zich meebrengt).

1.4 AI-plagiaat bewijzen: wat zegt de Raad van State?

In 2023 vielen in elk geval een twee uitspraken van de Raad van State op als het gaat over bewijs van fraude/plagiaat (al dan niet met ChatGPT). Uit die uitspraken kunnen we het volgende afleiden.

Niet voldoende voor bewijs: 'redelijk vermoeden'

Uit de [uitspraak](#) van 2 augustus 2023 volgt dat een ‘redelijk vermoeden’ van fraude niet voldoende is om een sanctie op te leggen. De aangedragen feiten en omstandigheden kunnen wel in de richting wijzen van fraude, maar als een andere toedracht niet kan worden uitgesloten, dan is de examencommissie niet in haar bewijslast geslaagd.

Wel voldoende voor bewijs: ‘buiten redelijke twijfel’

De [uitspraak](#) van 26 juli 2023 biedt handvatten om te komen tot een goed onderbouwde conclusie dat sprake is van AI-plagiaat. De Raad van State heeft in deze zaak geoordeeld dat er voldoende reden was om te twijfelen aan de vraag of het ingeleverde werk van de student door de student zelf was geproduceerd. Daarbij was cruciaal welke stappen er waren gezet in het vergaren van bewijs. Gedegen en gedocumenteerd onderzoek, niet alleen door de examencommissie, maar in eerste instantie ook door de vakdocent (examinator) is van groot belang om AI-plagiaat te bewijzen (en vervolgens te kunnen sanctioneren).

1.5 AI-plagiaat bewijzen: een rol voor de examinator

Het bepalen van de uitslag van een toets is de exclusieve bevoegdheid van de examinator (artikel 20 [Reglement examencommissies](#)). De examinator is verantwoordelijk voor het beoordelen van het werk en is ook verplicht om vermoedens van plagiaat aan de te melden aan examencommissie van de betreffende opleiding (hoofdstuk 7 [OER 23-24](#)).

In het geval van een vermoeden van AI-plagiaat wordt, met het oog op jurisprudentie, de volgende werkwijze aangeraden. Juist vanwege de complexiteit van het bewijzen van AI-plagiaat is het van belang dat (in afwijking van de werkwijze bij ‘reguliere’ plagiaat, waarbij de *examencommissie* het onderzoek doet naar het vermoeden van plagiaat) de examinator vanuit zijn expertise de volgende stappen zet. Dit is nodig om er voor te zorgen dat de examencommissie in een later stadium het vermoeden van AI-plagiaat voldoende kan onderbouwen. Het is van belang om deze werkwijze te communiceren met de student in de modulebeschrijving.

1. Volg de ontwikkeling van de student gedurende het vak. Laat de student voortuitgang zien en is die vooruitgang passend bij de kwaliteit van het eindproduct?
2. Als er twijfels ontstaan of het ingeleverde werk (in belangrijke mate) door de student zelf is geschreven, dan wordt de examinator aangeraden de volgende stappen te zetten:
 - a. Het ingeleverde werk kan worden vergeleken met eerder werk van de student als dat beschikbaar is (taalgebruik, complexiteit zinnen, stijl, etc.) en met eerdere beoordelingen van de student. Wekt het niveau en de inhoud van het werk van de student bevreemding, gelet op eerdere prestaties? Doe ook een check op de aangehaalde bronnen.
 - b. Toets als examinator het vermoeden bij een collega-examinator. Wijkt het stuk inderdaad af op basis van eerder ingeleverd werk van de student, ervaring bij begeleiding tijdens het onderwijs en zitten er opmerkelijke fouten in?
 - c. Als de collega-examinator het vermoeden deelt, ga dan als examinator met de student het gesprek aan over het ingeleverde werk en bespreek de bevindingen (bijvoorbeeld het onjuist gebruik van bronnen). Kan de student in deze mondeling toelichting de vragen over het werk beantwoorden, verbanden in de tekst toelichten? Kan de student ‘authentieke’ kennis, inzicht en vaardigheden laten zien? *Eventueel kan een tweede gesprek worden gevoerd in aanwezigheid van de tweede examinator.*
 - d. Maak een korte samenvatting van het/de gesprek(ken): welke vragen zijn gesteld, welke antwoorden zijn gegeven.

Ter overweging: De Raad van State oordeelde in deze uitspraak over deze specifieke zaak dat ook het aanbieden aan de student om een korte aanvullende opdracht te maken, een goede manier om te onderzoeken of de student (zelf) beschikt over de vereiste kennis en vaardigheden. Als de student afziet van een dergelijk aanbod en deze kans onbenut laat, kan daaraan de conclusie worden verbonden dat de twijfel over de kennis en kunde van de student niet is weggenomen.

Nota bene: Een dergelijke interventie raakt aan de haalbaarheid van toetsing, omdat het een noemenswaardige tijdsinvestering vraagt van studenten en docenten. Opleidingen kunnen hier zelf

een afweging in maken: is een dergelijke aanvullende korte opdracht een passende interventie, ook gelet op de omstandigheden van dit specifieke geval? Als opleidingen dit een passende interventie vinden, dan is het nodig dat dit vooraf in de modulebeschrijving als mogelijke interventie wordt opgenomen, zodat de student hiermee bekend is.

3. Als op basis van de bevindingen in stap 2 het vermoeden bij de examinerator dat sprake is van AI-plagiaat niet is weggenomen, meld dit dan (laagdrempelig) bij de examencommissie. Ook als er twijfel bestaat of er voldoende bewijs is. Zorg daarbij voor een onderbouwing van het vermoeden:
 - a. screenshots van fragmenten waar gebruik van AI wordt vermoed;
 - b. eventuele rapportage van een goedgekeurde AI-detectietool¹;
 - c. het ingeleverde werk van de student;
 - d. de schriftelijke samenvatting van het gesprek dat is gevoerd met de student.Stel de student op de hoogte van het feit dat het werk wordt voorgelegd aan de examencommissie vanwege een vermoeden van onregelmatigheden.

De examencommissie stelt vervolgens een onderzoek in.



In de praktijk kan het voorkomen dat de examinerator afziet van het melden van het vermoeden van een onregelmatigheid bij de examencommissie omdat hij twijfelt of sprake is van ongeoorloofd gebruik van (G)AI. Dit kan gevolgen hebben voor de waarde van het getuigschrift van de student en ook voor de kwaliteit van het toetsprogramma. Daarom is het aan te raden dat examencommissies zo veel als mogelijk stimuleren dat examineratoren een vermoeden van fraude laagdrempelig kunnen melden of hier advies over kunnen inwinnen.

1.6 AI-plagiaat bewijzen: de taak van de examencommissie [stappenplan]

Het is op grond van de WHW de taak van examencommissies om vast te stellen of inderdaad sprake is van AI-plagiaat en om de student die wordt verdacht van AI-plagiaat in de gelegenheid te stellen te worden gehoord. Vanwege de complexiteit van dit bewijs is het van belang dat de betrokken examinerator(en) (zoals hierboven beschreven) een (inhoudelijke) rol speelt in de beoordeling van het vermoeden van plagiaat.

Bij het bewijzen van fraude gaat het altijd om de weging van *feiten en omstandigheden die de conclusie kunnen dragen dat buiten redelijk twijfel is komen vast te staan dat de student plagiaat heeft gepleegd*. Om die reden is van belang om die feiten en omstandigheden goed te beschrijven, te wegen en te documenteren en een stappenplan te hanteren dat daarop aansluit.

Een algemeen uitgangspunt is de onschuldpresumptie van de student: de student mag niet bij voorbaat als schuldig worden gezien en moet een redelijk kans krijgen om het tegen hem bestaande bewijs te weerleggen.

Stappenplan 'buiten redelijke twijfel'

De examencommissie stelt het volgende onderzoek in:

1. Voer indien gewenst een (aanvullend) gesprek met de betreffende examinerator(en);
2. Voer een gesprek met de student, middels een hoorzitting. Doel van dit gesprek is om na te gaan of sprake is geweest van ongeoorloofd gebruik van (G)AI. Dit kan door vragen te stellen over de totstandkoming van het product en de inhoud ervan.
3. Maak voldoende gedetailleerde notulen van deze gesprekken (zie ook hieronder onder paragraaf 7).
4. Als de twijfel na stap 1 en 2 niet is weggenomen mag de examencommissie de conclusie trekken dat de fraude buiten redelijke twijfel is vast komen te staan, omdat de student de essentiële elementen van het ingeleverde werk niet heeft kunnen toelichten en niet over de kennis bleek te beschikken.

¹ Op het moment van schrijven van deze handreiking is er nog geen goedgekeurde tool beschikbaar.

5. Het nader technisch onderzoek (voor zover mogelijk) naar de digitale herkomstkenmerken van het schriftelijk werk door de examencommissie kan dan achterwege blijven.
6. Aan de student kan door de examencommissie een sanctie worden opgelegd. Verwijs daarbij naar de OER, specifieke onderdelen van het werk van de student, en voeg ook de verslagen van de gesprekken bij als bijlage. Voeg uiteraard een rechtsmiddelenclausule toe aan het besluit, waaruit volgt op welke manier en binnen welke termijn de student beroep kan aantekenen.



Als de examencommissie het bewijs niet rond krijgt maar tegelijk een zeer sterk vermoeden houdt van gepleegde fraude, dan kan dat wringen met de wettelijke taak van de examencommissie om de kwaliteit van toetsing te borgen. De Raad van State geeft de examencommissie in overweging om dan de [‘rode knop’](#) optie te gebruiken van de algemene ongeldigverklaring voor alle deelnemers van de toets. Bij de Haagse volgt die optie uit artikel 4.14 van OER 23-24. De [Raad van State](#) overweegt dat ‘onderzocht kan worden of in een geval als dit, waarin niet kan worden vastgesteld wat er precies is gebeurd, dergelijke maatregelen, die geen sancties zijn, kunnen worden genomen als de kennis en kunde van de studenten die de toets maakten niet meer objectief kan worden vastgesteld’.

Uiteraard is dit een vergaande maatregel, die overleg vereist met de faculteitsdirecteur. Het onderstreept te meer het belang van een (G)AI-resilient toetsprogramma.

1.7 Handvatten uit een uitspraak van het College van Beroep voor de Examens (CBE)

Bij het College van Beroep voor de Examens van De Haagse Hogeschool zijn nog maar weinig (G)AI gerelateerde zaken afgedaan. In een uitspraak van december 2023 was aan de orde dat een student werd verdacht van onrechtmatig gebruik van ChatGPT. In het besluit was door de examencommissie vermeld dat de examencommissie aan de student vragen had gesteld over de inhoud van het verslag en dat de student deze niet altijd kon beantwoorden. Ook kon de student het verband tussen bepaalde passages in het verslag niet uitleggen. Daarnaast heeft de examencommissie het ingeleverde verslag verder onderzocht en geconstateerd dat de schrijfstijl en het woordgebruik niet consistent zijn.

Er zijn uit deze uitspraak (in elk geval) twee lessen te trekken:

1. **Het verslag van de hoorzitting moet gedetailleerd zijn: welke vragen zijn gesteld en welke antwoorden zijn gegeven?**
In deze zaak ging het CBE voorbij aan de stellingen van de examencommissie omdat de examencommissie heeft nagelaten om een zodanig verslag op te stellen van de hoorzitting waaruit blijkt welke vragen zijn gesteld en welke antwoorden door de student zijn gegeven. Het CBE benadrukt dat van de examencommissie een dusdanig gedetailleerd verslag mag worden gevergd, vanwege de ingrijpendheid van een besluit om een student te sanctioneren.
2. **Het vermoeden van AI-plagiaat moet voldoende worden geconcretiseerd. De uitspraak van het CBE gaat hier niet verder op in, maar er valt uit af te leiden dat de feitelijke onderbouwing van het vermoeden van plagiaat heel belangrijk is. Daar kunnen dus ook het verslag van het gesprek met de examinerator en het gedetailleerd verslag van de hoorzitting voor worden gebruikt.**
Het CBE concludeert dat observaties van de examinerator weliswaar een vermoeden van plagiaat wekken, maar dat er nog niet voldoende bewijs is, ook niet door de constatering van de examencommissie (in het besluit) dat sprake is van een inconsistente schrijfstijl en woordkeuze. De examencommissie heeft deze bevindingen - tegen de achtergrond van een verslag van de student van behoorlijke omvang – onvoldoende geconcretiseerd, aldus het CBE.

Door de student is aangevoerd dat van hem niet kan worden verwacht dat hij bij de examencommissie een helder antwoord zou kunnen geven op vragen over een verslag dat hij ‘een kwartaal’ eerder had ingeleverd. Het CBE heeft achterwege gelaten om dit argument van de student te beoordelen. Toch kan hieruit worden afgeleid dat, als daar aanleiding toe is, kort na de beoordeling door de examinerator het werk te bespreken met de student. Er is dan immers weinig tijd verstreken en de student kan zich niet beroepen op de omstandigheid van de verstreken tijd.



In het kader van de motivering waarom een door de examencommissie opgelegde sanctie passend is, kan het nuttig zijn om de student voorafgaand aan de toets een 'integriteitsverklaring' te laten tekenen. Een dergelijke verklaring maakt de student nog eens extra bewust wat er van hem wordt verwacht. Opleidingen kunnen desgewenst een eigen integriteitsverklaring opstellen en gebruiken, die passend is bij de onderwijspraktijk.

Ter inspiratie: twee voorbeelden die worden gebruikt binnen De Haagse Hogeschool op het voorblad van een verslag, scriptie of rapport:

Ik, [naam van de student] verklaar hierbij dat ik de enige auteur van dit rapport ben. Alle hulp die ik van externe bronnen heb ontvangen, is gedocumenteerd in het rapport en staat vermeld na de conclusie onder de kop '.....'. Dit verslag is exclusief door mij gemaakt voor het vak '.....' aan De Haagse Hogeschool.

Door mijn toets in te dienen in Remindo bevestig ik het volgende:

(...)

dat ik geen ongeoorloofde hulp heb ontvangen bij het voorbereiden of schrijven van deze opdracht;

dat ik geen ongeoorloofde bronnen heb geraadpleegd (waaronder ChatGPT) bij het invullen van deze opdracht.

2. Het borgen van de kwaliteit van toetsen en examens

2.1 Taken en verantwoordelijkheden

Het is aan de opleiding/faculteit om te zorgen voor de kwaliteit van toetsing. Binnen faculteiten en opleidingen is men dan ook aan de slag gegaan met onder andere de vraag hoe eindkwalificaties van opleidingen worden geborgd in relatie tot het gebruik van AI, hoe aansluiting kan worden gezocht bij wensen en verwachtingen vanuit het werkveld, hoe gezorgd kan worden voor integratie van AI in het toetsprogramma en wat dat betekent voor de toetsvormen en op welke termijn. Opleidingen dienen de gebruikte toetsen (en het toetsbeleid) met het oog op (G)AI kritisch te beoordelen en zo nodig te herontwerpen zodat in de leeruitkomsten de competenties van de student (zelf) worden getoetst. Zie hiervoor de documenten waar naar wordt verwezen in paragraaf 2.3. Belangrijke zaken, zoals het toetsbeleid op opleidingsniveau (in de context van (G)AI) kan in het Opleidingskader worden vastgelegd.

Om de borgende taak goed uit te kunnen voeren, is het van belang dat de examencommissie en de opleiding (waaronder management en examinatoren) het gesprek voeren over validiteit, betrouwbaarheid, functie en condities op verschillende niveaus van toetsing in relatie tot (G)AI. Een belangrijk uitgangspunt voor de borgende rol van de examencommissie is het toetsprogramma. Vanuit het toetsprogramma zijn vervolgens vragen te stellen over de toetsen. De [KITplus](#) is een mooi hulpmiddel en kan de examencommissie ondersteunen om in hun borgende bevoegdheid tot passende adviezen aan de opleidingen te komen.

2.2 Richtlijnen van de examencommissie

Deskundige examinatoren zijn een belangrijke waarborg voor de kwaliteit van toetsing. De aanwijzingsbevoegdheid van de examencommissie hangt nauw samen met de verantwoordelijkheid van de examencommissie voor het borgen van de kwaliteit van toetsen en examens. De aanwijzing van examinatoren is een instrument om de kwaliteit van toetsing binnen de opleiding te beheersen en dat instrument moet de examencommissie bewust moet toepassen.

De WHW geeft examencommissie de bevoegdheid om zelf richtlijnen en aanwijzingen vast te stellen. In de context van (G)AI kan de examencommissie dan bijvoorbeeld eisen stellen aan de AI-geletterdheid van examinatoren (als beoordelaars en ontwikkelaars van toetsen) of (aanvullende) richtlijnen voor examinatoren om AI-plagiaat te bestrijden, uiteraard binnen de hogeschoolbrede kaders.

2.3 Kwaliteitsprincipes toetsing en (G)AI

Het toetsbeleid van De Haagse Hogeschool ([De Haagse Toetsing](#)) geeft gemeenschappelijke kaders voor de wijze waarop we binnen De Haagse Hogeschool toetsen vormgeven en de kwaliteit van toetsen borgen. Met de voortschrijdende ontwikkelingen op het gebied van (G)AI, worden verschillende aspecten van ons huidige toetskader en de onderliggende kwaliteitsprincipes van toetsing in meer of mindere mate beïnvloed.

Sinds de lancering van ChatGPT in november 2022 en de daaropvolgende ontwikkelingen op het gebied van (G)AI en onderwijs zijn er binnen de Hogeschool verschillende documenten en handreikingen gepubliceerd. Op het gebied van toetsing is er bijvoorbeeld de [Toolkit AI in je toetsontwerp](#), bestaande uit:

- handreiking ChatGPT en toetsing voor docenten;
- praatplaat ChatGPT en toetsing;
- handreiking academische integriteit voor studenten;
- interactieve powerpoint AI in je toetsontwerp;
- kaartenset leren zichtbaar maken met o.a. suggesties voor toetsvormen.

In het huidige document zijn de uitgangspunten van de Haagse Toetsing die relevant zijn in de context van (G)AI verder uitgewerkt in specifieke uitgangspunten voor (G)AI en toetsing. Doel van deze

uitgangspunten is om voor de korte termijn kaders te bieden aan opleidingen om de kwaliteit van toetsen zo goed mogelijk te kunnen garanderen. De mogelijkheden en ontwikkelingen op het gebied van AI worden nauw gevolgd. Indien nodig worden de uitgangspunten hierop aangepast.

Gezien de grote diversiteit aan opleidingen, beroepsprofielen en student- en docentgroepen bieden deze uitgangspunten zo veel als mogelijk brede kaders voor opleidingen, examencommissies en examinatoren. Deze kaders kunnen opleidingen vervolgens in samenspraak met relevante belanghebbenden vertalen naar hun specifieke onderwijspraktijk, o.a. op basis van de facultaire visie op toetsing en de kaders uit het toetshandboek.

De uitgangspunten van de Haagse Toetsing in de specifieke context van (G)AI zijn als volgt:

a. Het gebruik van (G)AI in toetsing is in lijn met de beroepspraktijk

De ontwikkelingen op het gebied van (G)AI hebben in meer of mindere mate ook invloed op het toekomstige beroepenveld en daarmee de beroepsbekwaamheid van een student en de eindkwalificaties van een opleiding.

We dragen zorg voor authentieke toetsing die afgestemd is met de beroepspraktijk en aansluit op wat er van toekomstige professionals wordt gevraagd. Dit geldt niet alleen voor de gevraagde producten en handelingen en de gehanteerde beoordelingscriteria, maar ook voor de manier waarop (G)AI gebruikt kan en mag worden. Dit betekent dat de manier waarop (G)AI gebruikt wordt in de beroepspraktijk ook als zodanig terugkomt in de toetsing en beoordeling.

b. We zijn transparant als het gaat om het gebruik van (G)AI in toetsing

Opleidingen communiceren duidelijk op welke manier het gebruik van (G)AI bij toetsing door studenten wel/niet is toegestaan en in welke mate, wat mogelijke consequenties zijn bij het ongeoorloofd gebruik van (G)AI. Opleidingen begeleiden studenten bij verantwoord gebruik van (G)AI in onderwijs, toetsing en in hun toekomstige beroep.

Het is niet toegestaan dat studenten persoonsgegevens inbrengen in (G)AI tools. Persoonsgegevens zijn alle gegevens die iets zeggen over een specifiek persoon en waaruit je kunt afleiden om wie het gaat. Ook losse gegevens die samengevoegd kunnen worden en daarmee iets zeggen over een bepaald persoon, zijn persoonsgegevens. Voorbeelden zijn voornaam, achternaam, contactgegevens, telefoonnummer, foto's, interviews, gezondheidsgegevens, locatiegegevens en nog veel meer. Studenten moeten duidelijk aangeven of en hoe zij (G)AI hebben gebruikt bij hun toetsen (attributie), zoals vereist in de onderwijs- en examenregeling en beschreven in de handreiking voor academische integriteit. Dit houdt in dat studenten inzichtelijk maken welke tool is gebruikt, welke prompt of instructies zijn gegeven, welke output zij hebben verkregen en hoe deze output al dan niet is verwerkt in de toets. Daarnaast geven studenten een bronvermelding in de lopende tekst en de bronnenlijst (zie bijvoorbeeld ook de APA-richtlijnen voor bronverwijzing van (G)AI).

c. Wij houden onze toetsing inclusief en toegankelijk door het gebruik van (G)AI-tools in toetsing niet verplicht te stellen voor studenten

De Haagse Hogeschool heeft (nog) geen officiële overeenkomsten met aanbieders van (G)AI. Bovendien zijn duurdere (G)AI-tools doorgaans beter dan gratis (G)AI-tools, wat ongelijkheid creëert tussen studenten onderling. Dit betekent dat, tot er officiële overeenkomsten en licenties zijn en studenten gelijke toegang hebben tot (G)AI hulpmiddelen, we het gebruik van (G)AI in toetsing door studenten niet verplicht stellen en dat we - in lijn met de leeruitkomsten - alternatieven bieden voor studenten die geen gebruik willen en/of kunnen maken van (G)AI. Natuurlijk kan het gebruik van (G)AI ook bijdragen aan inclusievere toetsing. Dit valt op dit moment echter buiten de scope van het huidige document. Wanneer De Haagse Hogeschool wel de bedoelde licenties heeft verkregen, zal worden bezien of en hoe deze kunnen worden ingezet bij toetsing.

d. We laten het werk van studenten alleen beoordelen door aangestelde examinatoren

Alleen examinatoren die zijn aangesteld door de examencommissie zijn bevoegd om het werk van studenten te beoordelen. (G)AI is dat niet. Examinatoren dragen in lijn met de uitgangspunten van de Haagse Toetsing ook zorg voor het geven van kwalitatieve, constructieve en persoonlijke feedback (en

feedforward). Het uitgangspunt is dat de kwaliteit van door (G)AI gegenereerde feedback vooralsnog niet toereikend is. De rol van de examiner blijft onverminderd belangrijk in het formuleren van feedback en feedforward aansluitend op de specifieke context en leerbehoeften van de student en het vertalen van observaties naar een (summatieve) beoordeling.

Bovendien is het uploaden van het werk van studenten in (G)AI zonder expliciete toestemming van de student in strijd met de [Privacyverklaring](#) van De Haagse Hogeschool. Op dit moment lopen er pilots naar de mogelijkheden voor (G)AI om ondersteuning te bieden in het feedback- en beoordelingsproces. Deze pilots worden in een veilige omgeving afgenomen en vanuit verschillende expertisegebieden gemonitord en geëvalueerd vanuit de AI Hub van De Haagse Hogeschool, met raadpleging van verschillende experts. Indien nodig worden de uitgangspunten hierop aangepast.

e. We waarborgen de kwaliteit van toetsing in de context van AI

De volgende uitgangspunten zijn daarbij van belang. Deze uitgangspunten zijn gebaseerd op de onderliggende kwaliteitsprincipes van toetsing, die met het oog op (G)AI niet altijd anders zijn, maar die wel meegenomen moeten worden in het toetsontwerp.

1. **We gaan er in het toetsontwerp vanuit dat studenten bij take-home toetsen gebruik maken van alle middelen die zij tot hun beschikking hebben.** Dit betekent niet dat ongeoorloofd gebruik van (G)AI of welke ander bron dan ook is toegestaan, maar dat bij toetsontwerp interventies worden gedaan om te voorkomen dat studenten ongeoorloofd gebruik kunnen maken van (G)AI. Je hebt namelijk geen zicht op het gebruik van hulpmiddelen en detectie van (G)AI schiet tekort. Een take-home toets dient daarom zo ontworpen te worden dat de student eigen kennis, vaardigheden en competenties laat zien, ook als zij alle hulpmiddelen tot hun beschikking hebben. Dit kan mogelijk betekenen dat een toetsopdracht wordt aangepast, evenals de beoordelingscriteria en prestatieindicatoren (voor suggesties zie toolkit AI in je toetsontwerp).
2. **We ontwerpen toetsen vanuit heldere en duidelijke leeruitkomsten die o.a. herkenbaar zijn voor de beroepspraktijk (zie ook handleiding formuleren leeruitkomsten).**
3. **We zorgen voor afstemming tussen leeruitkomsten, toetsing en beoordeling.** Als studenten AI mogen gebruiken, passen we waar nodig de leeruitkomsten en de daarvan afgeleide beoordelingscriteria en indicatoren aan. Hiermee borgen wij dat de prestatie die studenten gevraagd worden te leveren dekkend blijft voor de leeruitkomsten en beoordeeld kan worden op basis van de vastgestelde indicatoren. Hierbij kun je denken aan:
 - het niveau: is het nodig om aanpassingen te maken in het cognitieve niveau, de complexiteit of mate van autonomie met het oog op (G)AI gebruik?
 - de weging: is de weging van de beoordelingscriteria en/of leeruitkomsten nog steeds representatief voor de prestatie die wordt geleverd? Is het bijvoorbeeld nodig om een mondelinge component zwaarder te wegen dan een schriftelijke component?
 - de beoordelingscriteria: op welke aspecten beoordeel je de prestatie? Is het nodig om op andere criteria te beoordelen als een student (G)AI mag gebruiken in de toetsing?
4. **Waar mogelijk en passend kiezen we voor directe bewijzen en 'live' assessments.** 'Live' assessments zijn toetsen waarin studenten actief moeten handelen (denk aan presentaties, criterium-gerichte interviews, practica en performance-assessments). Directe bewijzen zijn eigen observeerbare handelingen en/of producten. Indirecte bewijzen niet direct observeerbaar, zoals een reflectie of feedback van anderen. Let op: als reflecterend vermogen onderdeel is van de leeruitkomsten dan kan een (schriftelijke) reflectie wel gezien worden als een direct bewijs. In dat geval zou je er voor kunnen kiezen om een reflectie mondeling te toetsen.
5. **Studenten blijven volledig verantwoordelijk voor de inhoud van hun werk** (ook in het geval van groepswork) en dragen zorg voor het feit dat het ingeleverde werk de eigen kennis, vaardigheden, competenties in relatie tot de leeruitkomsten weerspiegelt (attributie). Afhankelijk van o.a. de leeruitkomsten en daarvan afgeleide beoordelingscriteria wordt bepaald op welke wijze en in welke omvang (G)AI gebruik al dan niet is toegestaan. Gegenereerde output van (G)AI dient door de student altijd gecheckt te worden, bijvoorbeeld op basis van geverifieerde bronnen.
6. **Examinatoren maken eenduidige afspraken over de plek van (G)AI in het toetsontwerp om de consistentie en betrouwbaarheid in het toets- en beoordelingsproces te borgen.** Een toetsresultaat mag niet afhankelijk zijn van irrelevante omstandigheden, bijvoorbeeld de strengheid van een bepaalde docent, de interpretatie van beoordelingscriteria door verschillende docenten of omstandigheden tijdens de toetsafname. Het is daarom belangrijk voor examinatoren om hier

duidelijke afspraken over te maken en vast te leggen aansluitend op de uitgangspunten in dit document.

7. **Toetsen zijn haalbaar en uitvoerbaar voor zowel examinatoren als studenten.** Dit houdt in dat de student voldoende tijd en middelen moet hebben om de toets voor te bereiden en te maken. De examinerator moet voldoende tijd hebben om de toets te ontwerpen, af te nemen en te beoordelen. Extra interventies ten gevolge van (G)AI ontwikkelingen (bijvoorbeeld een herontwerp van een toets, extra nakijktijd, en/of andere kwaliteitsborgende activiteiten van examinerator en/of examencommissie) moeten gefaciliteerd worden door de opleiding.
8. **We blijven de kwaliteit van onze toetsing evalueren.** Toetsanalyses (zowel kwantitatief als kwalitatief) zijn een vast onderdeel van de toetscyclus. In de context van (G)AI is het belangrijk om extra alert te zijn op de validiteit, betrouwbaarheid, transparantie, inclusiviteit en of de toetsing passend is bij ontwikkelingen in het beroepenveld (authenticiteit).

3. Casuïstiek uit de praktijk

Onderstaande casussen zijn opgehaald vanuit verschillende opleidingen binnen De Haagse Hogeschool. Per casus geven we een toelichting en suggesties hoe examencommissies en opleiding met een dergelijke casus om zouden kunnen gaan, vanuit de bestaande kaders binnen de hogeschool.



Een docent heeft een handige manier gevonden om een toets met open vragen te laten nakijken door ChatGPT. De student klaagt hierover bij de examinerator omdat de student van mening is dat het ten koste gaat van de kwaliteit van de beoordeling. De examencommissie was niet bekend met deze manier van nakijken door de docent.

Ook hier geldt: het werk van een student mag niet zonder toestemming van de student worden ingevoerd in een (G)AI tool. Hieronder bespreken we nog een aantal aanvullende aspecten, waarbij we meer inzoomen op de taak en rol van de examinerator, ook in relatie tot de examencommissie.

Uitgangspunt is dat het werk van studenten wordt beoordeeld door een - door de examencommissie aangewezen - examinerator. Het is de exclusieve bevoegdheid van de examinerator om de uitslag van de toets vast te stellen. (G)AI kan uiteraard niet worden beschouwd als een examinerator. Het is nadrukkelijk niet toegestaan om zonder menselijke controle het werk van een student te beoordelen. De door de examencommissie aangewezen examinerator blijft altijd eindverantwoordelijk voor de beoordeling. De student heeft recht op 'rijke' feedback van een docent die voldoende bedreven is (en blijft) in het (zelf) geven van die feedback.

Van de examinerator mag worden verwacht dat hij aantoonbaar gekwalificeerd is op het gebied van toetsing en dat hij zich ervan bewust is dat de privacy en het intellectueel eigendom van de student niet zijn gewaarborgd als een opdracht van een student wordt ingevoerd in een applicatie als ChatGPT. De examencommissie borgt de kwaliteit van toetsen en examens. Daar hoort ook bij dat de examencommissie zich verdiept in de vraag hoe een beoordeling van een toets tot stand is gekomen. De examencommissie kan de examinerator vragen hoe hij de toets heeft beoordeeld om na te gaan of de examinerator het toetsbeleid zoals vastgelegd in de OER en de richtlijnen en aanwijzingen in acht genomen heeft. Als dat niet het geval is, dient het gesprek daarover te worden gevoerd met de betrokken examinerator.

Zoals beschreven in hoofdstuk 2.3 onder d van deze handreiking lopen op dit moment pilots om de mogelijkheden te onderzoeken voor ondersteuning door (G)AI in het feedback- en beoordelingsproces. Deze pilots worden in een veilige omgeving afgenomen en vanuit verschillende expertisegebieden gemonitord en geëvalueerd, ook met het oog op de EU AI Act (recital 35), waar het gebruik van AI-systemen voor het evalueren van leerresultaten wordt geclassificeerd als hoog risicovol. Als de uitkomst van pilots hier aanleiding toe geven, worden de uitgangspunten aangepast.

Het is om bovengenoemde redenen van belang dat examineratoren niet zonder inmenging van de examencommissie experimenteren met het inzetten van (G)AI bij het beoordelen van werk van studenten.



Een docent heeft het gebruik van AI geïntegreerd in de toetsing. De beoordelingscriteria en leeruitkomsten zijn hetzelfde gebleven. De examencommissie is van mening dat de prestatie die gevraagd wordt van de student niet meer voldoende in lijn is met de beoordelingscriteria en indicatoren en niet meer voldoende dekkend is voor de leeruitkomsten.

De examencommissie oefent haar bevoegdheid uit door te borgen dat de leeruitkomsten en eindniveau door de individuele student worden behaald. Ook ziet de examencommissie er op toe dat het toetsprogramma in totaliteit de vastgestelde opleidingskwalificaties dekt. In een kwestie als deze is mogelijk de borging van het eindniveau in het geding.

De toetsmatrijs moet inzichtelijk maken hoe leeruitkomsten, eindkwalificaties en toetsing zich tot elkaar verhouden. Is de toetsmatrijs in lijn met de gekozen toetsvorm? Bijvoorbeeld ten aanzien van de complexiteit van de taak of het vereiste cognitieve niveau, of het op te leveren product dat kan veranderen op het moment dat bij de toets gebruik mag worden gemaakt van (G)AI.

In een gesprek van de examencommissie/toetscommissie met de examinerator zou de vraag aan bod moeten komen of de toetsvorm nog past bij de te behalen leeruitkomsten. Is het meerogenprincipe toegepast bij de constructie van de toets en heeft het gesprek over de gemaakte keuzes met de collega examinerator plaatsgevonden? Wat is daar de opbrengst van en welke afwegingen zijn gemaakt? Zo nodig moet het toetsontwerp worden aangepast met het oog op de (G)AI ontwikkelingen.

Als de toets al is afgenomen op het moment dat de examencommissie tot deze conclusie komt, is het denkbaar dat een aanvullende opdracht wordt afgenomen waarin de ontbrekende onderdelen nog worden afgetoetst. In een uiterst geval kan de toets ongeldig worden verklaard, maar dit betreft een ingrijpende beslissing met vergaande gevolgen. In elk geval moet het gesprek (zoals hierboven beschreven) hierover worden gevoerd en dient in het kader van de PDCA cyclus het toetsontwerp te worden gewogen en zo nodig aangepast.



Een docent moedigt studenten aan om AI te gebruiken bij het maken van hun toets. Op voorwaarde dat studenten transparant zijn over hun gebruik van AI. Dezelfde docent dient een zaak in bij de examencommissie: een student heeft veel te veel gebruik gemaakt van AI.

Het is van belang dat helder aan de student wordt gecommuniceerd in welke mate en op welke manier de student gebruik mag maken van AI bij het maken van de toets (zie paragraaf 1.2) en dat het belang van een deugdelijke bronvermelding wordt onderstreept. Het moet de student duidelijk zijn dat hij werk dient in te leveren dat zijn leeruitkomsten en competenties weerspiegelt.

Het is de verantwoordelijkheid van de opleiding/de examinerator om er zorg voor te dragen dat de toets van goede kwaliteit is en voorzien van heldere beoordelingsprocedures en -criteria. Bovendien moet de toets samengesteld zijn in overeenstemming met de OER, de studiegids en informatie uit de modulebeschrijving

Als het ingeleverde werk tekortschiet (zoals in dit voorbeeld) is het uiteraard niet zo dat daarmee altijd sprake is van fraude. Als uit de bronvermelding blijkt dat er teveel gebruik gemaakt is gemaakt van (G)AI, kan daarover het gesprek worden gevoerd met de student. Het is aan de docent om aan de hand van beoordelingscriteria de toets te beoordelen en eventueel een onvoldoende beoordeling te geven.

Tot slot: Een student kan niet worden *verplicht* om (G)AI te gebruiken bij het maken van een toets met het oog op de privacy. Op dit moment is niet voldoende duidelijk wat er met geüploade stukken in open source bots gebeurt.

4. Bronnen

Deze handleiding kwam tot stand op basis van een brede oriëntatie op AI en de borging van toetskwaliteit. Ook zijn gesprekken gevoerd binnen en buiten de hogeschool met betrokkenen bij het onderwerp.

Ter inspiratie zijn bij het schrijven van deze handleiding verschillende publicaties van andere hoger onderwijsinstellingen en een aantal andere bronnen geraadpleegd.

[EU AI Act](#)

Saxion Hogeschool (2023). *Handleiking ChatGPT voor studenten*.
Fontys Hogeschool (2023). *Generative Artificial Intelligence (GAI). Experimenteer maar blijf alert*.
Hogeschool Utrecht (2023). *Handleiking Generatieve AI en toetsing*.
Handleiking Hogeschool InHolland, 2023.
Universiteit Leiden (2023). *AI in het onderwijs*.
Hogeschool InHolland (2023). *Memo fraude rondom ChatGPT*.
Hogeschool Windesheim (2023). *Generatieve AI op hogeschool Windesheim*.
HAN University of Applied Sciences (2023). *Handleiking ChatGPT&Toetsing*

Slimmer onderwijs met AI, 2023. Npuls, [Slimmer-onderwijs-met-AI-Npuls.pdf](#)

KIT plus. *Instrument voor het borgen van de kwaliteit van toetsen, toetsprogramma's en toetsbeleid* (2018)

Vereniging Hogescholen (2021). [Handleiking Examencommissies](#).

Last, B. en Sprakel, T. (2023). Chatten met Napoleon. Werken met generatieve AI in het onderwijs. Boom.

[Aanbiedingsbrief \(overheidsbrede\) visie op generatieve AI, d.d 18 januari 2024.](#)

[Toezicht op AI&Algoritmes. Autoriteit Persoonsgegevens.](#)