

AI-BESTENDIG TOETSEN

Een toolkit voor het analyseren van toetsen



Themagroep AI - Toetsing van HAN Leernetwerk Toetsing
10 oktober 2024

OPEN UP NEW HORIZONS.

HAN_UNIVERSITY
OF APPLIED SCIENCES

COLOFON

VERSIE 1

DIT DOCUMENT IS TOT STAND GEKOMEN MET BIJDAGEN VAN:

Naam	Rol
Sabrina Bloemen	Toetsdeskundige en onderwijskundig adviseur Services Onderwijs, Onderzoek en Kwaliteitszorg
Jessica Koning	Toetsdeskundige Academie Business en Communicatie
Minke Zeillemaker	Onderwijskundige Academie Business en Communicatie
Mirjam Raanhuis	Onderwijskundige Academie Mens en Maatschappij
Jochem ten Böhmer	Onderwijskundige blended leren Services Onderwijs, Onderzoek en Kwaliteitszorg
Jorn Bunk	Onderwijskundige en AI-expert Services Onderwijs, Onderzoek en Kwaliteitszorg
Martijn Hendriks	Docent Academie Rechten en Academie IT & Mediadesign, toetscommissielid
Maartje Gijssels-Keijsers	Onderwijskundig adviseur Academie Organisatie en Ontwikkeling
Sara Struik	Onderwijskundige Academie Organisatie en Ontwikkeling
Dirk Weijers	Toetsdeskundige Academie Gezondheid en Vitaliteit
Frank Vonk	Beleidsmedewerker Kwaliteitszorg onderwijs en onderzoek
Marjoleine Dobbelaer	Onderzoeker Kwaliteit van Leraren
Marion van de Wijdeven	Onderwijsadviseur en trainer onderzoeksteam Kwaliteit van Leraren
Edwin Buijs	Onderzoeker Kwaliteit van Leraren
Mariëlle Wisselink	Onderwijskundig beleidsmedewerker International School of Business
Rob Warmershoven	Hoofddocent International School of Business
Ingeborg van den Heuvel	Docent International School of Business
Wout Brink	Docent Ondernemerschap & Retailmanagement

Dit document is vastgesteld op 10 oktober 2024 door de themagroep AI & Toetsing van het HAN Leernetwerk Toetsing.

Op deze uitgave is de Creative Commons Naamsvermelding 4.0-licentie van toepassing. Maak bij gebruik van dit werk vermelding van de volgende referentie: Themagroep AI – Toetsing van HAN Leernetwerk Toetsing (2024). Risicoanalyse Toetsvormen en AI. Nijmegen/Arnhem: HAN



INHOUDSOPGAVE

1	INLEIDING.....	5
2	FLOWCHART AI-BESTENDIG TOETSEN.....	7
	BIJLAGE A. ONDERBOUWING TOETSVORMEN	8
A.1	AI-BESTENDIG	8
A.2	MIDDELHOOG RISICO.....	9
A.3	HOOG RISICO	10
	BIJLAGE B. TOETSACTIVITEITEN	12
	BIJLAGE C. LIJST MET MAATREGELEN.....	13

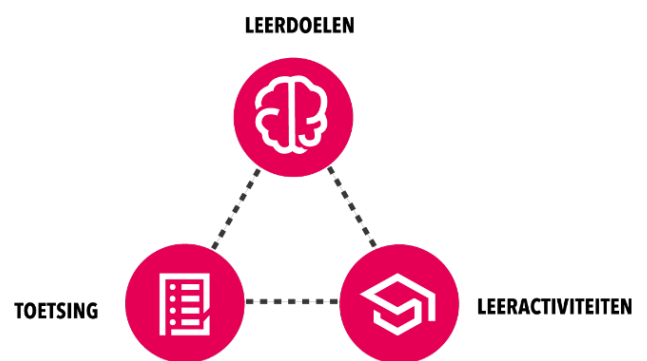
1 INLEIDING

AI-systemen, en specifiek generatieve AI-systemen, hebben een duidelijk effect op hoe studenten leren, wat ze (moeten) leren en hoe opleidingen toetsen of studenten de gewenste leeruitkomsten beheersen. Dit vraagt van opleidingen om kritisch te kijken naar hun onderwijs en toetsing en aanpassingen te maken waar nodig. Deze toolkit is een handreiking om de toetsen van een opleiding tegen het licht te houden en te kijken in hoeverre ze AI-bestendig zijn.

AI-bestendig onderwijs: Onderwijs met constructive alignment, waarin bij gebruik van AI dit enkel ingezet wordt ter ondersteuning van het leerproces en/of de toetsing.

AI-bestendig onderwijs kan zowel zijn met AI als zonder AI en betreft het gebruik van AI die wel en die niet is toegestaan. Effectief onderwijs is gegrond in een stevige *constructive alignment*, waarbij de leeruitkomsten, onderwijsactiviteiten en toetsing een logisch consistent geheel vormen zowel inhoudelijk als qua vorm. Voor AI-bestendig onderwijs houdt dit bijvoorbeeld in dat als volgens de leeruitkomst een student zelf over schrijfvaardigheden moet bezitten, het gebruik van AI in de toets niet gewenst is, terwijl dat bij onderwijs waarin andere leeruitkomsten centraal staan geen probleem hoeft te zijn.

CONSTRUCTIVE ALIGNMENT



Figuur 1: Constructive Alignment

Let op:

1. Het doel van deze toolkit is zicht krijgen op de AI-bestendigheid van de toetsen in de opleiding. De uitkomsten zeggen daarmee nog niets over de validiteit en betrouwbaarheid van de toetsen.
2. AI is een veld dat zich snel ontwikkelt. Daarom doen we een appèl op opleidingen om kritisch te kijken en te blijven kijken naar hun toetsing in het licht van het gehele toetsprogramma. Neem daarbij de onderliggende onderbouwing van de risico-indicatie en de te nemen maatregelen mee in de analyse.

1.1 Toelichting bij de toolkit

Dit document is een toolkit voor opleidingen om het gesprek over de AI-bestendigheid van de toetsing te voeren. De toolkit bestaat uit een flowchart van waaruit verwezen wordt naar onderliggende bijlagen om de verschillende stappen uit de flowchart nader te onderzoeken. Dit geeft opleidingen de

mogelijkheid om in kaart te brengen waar risico's in het toetsprogramma ten aanzien van AI-bestendigheid zitten en op basis van de uitkomst van de analyse prioriteiten te bepalen voor een (door)ontwikkeling van het onderwijs. Gebruik de toolkit bij voorkeur samen met examencommissie-, toetscommissie- en curriculumcommissieleden om het gehele toetsprogramma te analyseren en om vervolgcacties te bespreken. Daaronder valt ook de terugkoppeling aan het docententeam en het betrekken van docenten bij het doorvoeren van aanpassingen/verbeteringen. Ook kan de toolkit voor individuele toetsen met of door de desbetreffende docent(en)/examinator(en) gebruikt worden.

1.2 Aandachtspunten en tips voor het gesprek

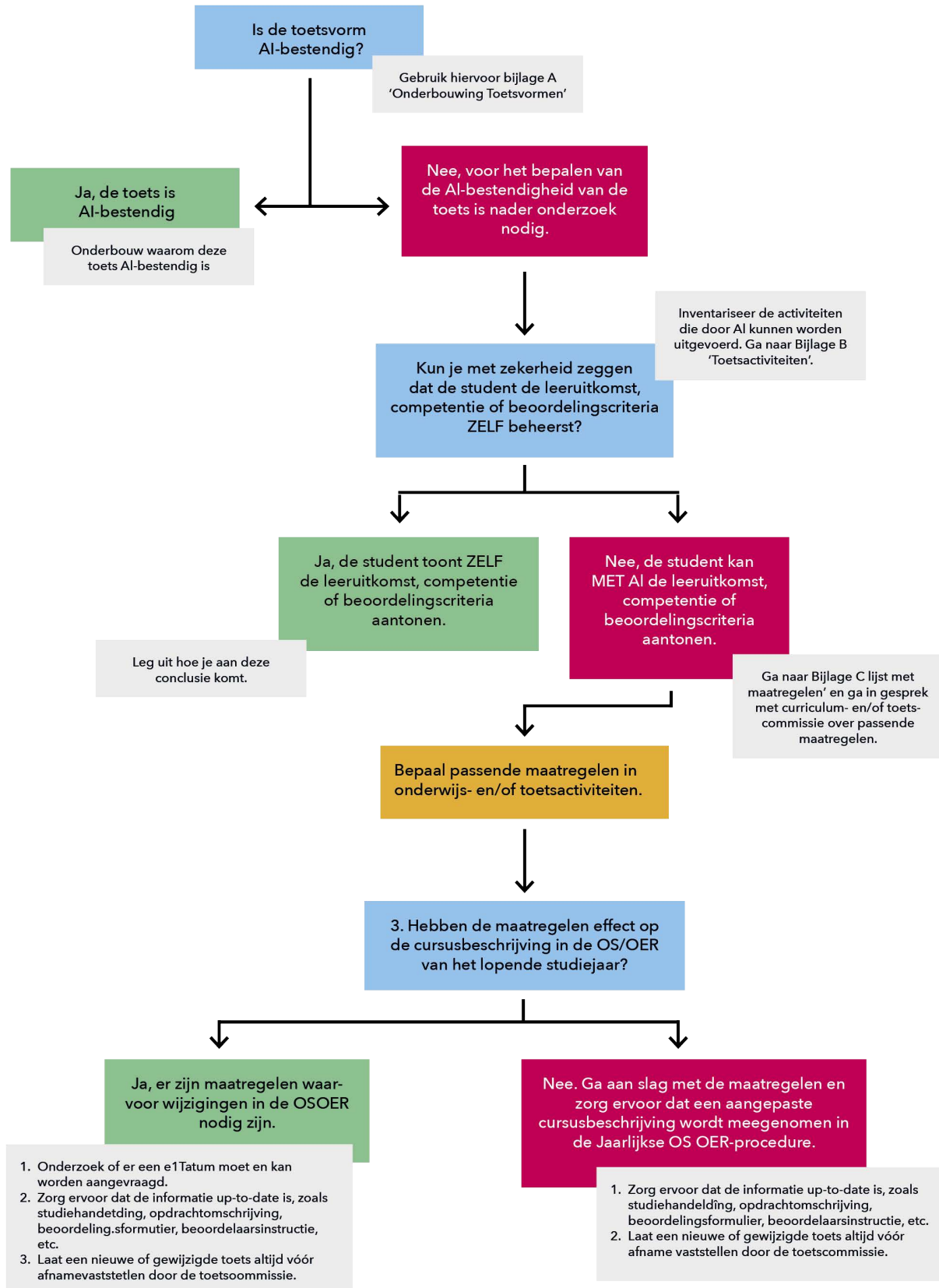
Bij het voeren van het gesprek waarbij de flowchart doorlopen wordt, geven we de volgende tips:

- Betrek in het gesprek verschillende relevante actoren uit de opleiding.
- Betrek in het gesprek de toets- of onderwijskundige(n) van je academie of neem contact op met het HAN Leernetwerk Toetsing. Het Teamskanaal: [HAN Leernetwerk Toetsing | Algemeen | Microsoft Teams](#)¹
- Het doel van het gesprek is niet om toetsvormen te verbieden, maar met elkaar op zoek te gaan naar de juiste, passende oplossingen in het geval van AI-fraudegevoelige toetsen. Omdat bij sommige toetsen de AI-bestendigheid afhangt van meer dan alleen de toetsvorm, is het zinvol om per toetsvorm tot een antwoord te komen dat start met “ja, mits” of “nee, tenzij” in plaats van een harde “ja” of “nee”. Dit zorgt voor het maken van bewuste keuzes en het bieden van handelingsperspectief.
- Leidend bij het analyseren van de toetsvormen is wat de examiner(en) écht wil(len) zien van de student, niet de OSIRIS-code en/of de benaming van de toetsvorm (in dit document). Uit de praktijk blijkt namelijk dat het een niet hoeft samen te vallen met het ander. Zo kan als toetsvorm een portfoliotoets aangegeven zijn, terwijl in de praktijk de student dit portfolio ook moet onderbouwen en toelichten in een mondeling gesprek of CGI.
- Bekijk een toets niet als een losse entiteit. Toetsen zijn afgeleid van leeruitkomsten en vormen samen met leeractiviteiten het gehele curriculum. Ook hoeft een enkele middelhoog-risicotoets (zie bijlage A) geen probleem te zijn als de kwaliteit van het toetsprogramma als geheel hoog is en er binnen het onderwijs goed zicht is op het leerproces van de student.
- Kijk goed naar de onderbouwing per toetsvorm in deze risicoanalyse. Het daadwerkelijke risico van een toetsvorm is afhankelijk van meerdere factoren, zoals de manier waarop er in het onderwijs zicht is op de ontwikkeling van de student en de wijze van toetsafname.

¹<https://teams.microsoft.com/l/team/19%3Aa52986629b804e3db124edbf0bdee00e%40thread.tacv2/conversations?groupId=8ac74beb-bde3-4a2d-8f85-b03952141a0b&tenantId=5d73e7b7-b3e1-4d00-b303-056140b2a3b4>

2 FLOWCHART AI-BESTENDIG TOETSEN

Doorloop onderstaande stappen en maak gebruik van de bijlagen.



BIJLAGE A. ONDERBOUWING TOETSVORMEN

Grofweg kan al een eerste indeling worden gemaakt tussen toetsvormen die wel (AI-fraudegevoelig) of geen (AI-bestendig) risico vormen:

AI-fraudegevoelige toetsvormen: Alle toetsproducten, die zonder direct toezicht van een examiner of surveillant zijn geproduceerd en waarvan één of meerdere leeruitkomsten zonder gebruik van AI moeten worden aangetoond. Deze vormen een middelhoog tot hoog risico op fraude met AI.

AI-bestendige toetsvormen: Alle toetsproducten, die onder direct toezicht van een examiner of surveillant is geproduceerd **of** afgenomen, en toetsing waarbij het gebruik van AI-onderdeel is (geworden) van de leeruitkomst of niet problematisch is voor het aantonen van de leeruitkomsten.

De risicoanalyse hieronder geeft drie categorieën om meer ruimte te geven aan de context van de toets(vorm). We herhalen hier graag dat de oplossing voor AI-fraudegevoeligheid niet is om alleen nog AI-bestendige toetsvormen aan te bieden. De oplossingen zijn te vinden in de gehele constructieve alignment, zoals te zien bij de maatregelen in Bijlage C.

AI-BESTENDIG



- Beroepshandeling (met live observatie)
- Geroosterde schriftelijke inzichttoets
- Geroosterde schriftelijke kennistoets
- Geroosterde digitale inzichttoets
- Geroosterde digitale kennistoets
- Practicum live observatie
- Beroepsproduct (fysiek inleveren)
- Criterium Gericht Interview
- Practicum opname live met aanwezigheid docent
- Mondelinge toets

MIDDELHOOG RISICO



- Beroepshandeling (digitaal inleveren, opname)
- Presentatie
- Practicum opname, zonder aanwezigheid docent
- Beroepsproduct (fysiek inleveren)
- Audiovisuele opdracht

HOOG RISICO



- Reflectieverslag
- Beoordelingsportfolio
- Beroepsproduct (digitaal inleveren)
- Schriftelijke opdracht met tijdslimiet
- Verantwoordingsverslag
- Schriftelijke opdracht

A.1 AI-BESTENDIG

Beroepshandeling (met live observatie)

Als de beroepshandeling wordt geobserveerd, beoordeelt de examiner wat hij/zij ziet. Ook als bij de beroepshandeling gebruik gemaakt wordt van een computer en/of AI, dan ziet de examiner wat er gebeurt en kan het gebruik van AI-onderdeel zijn van de beoordeling.

Mondelinge toets

De examiner(en) voert het gesprek live met een kandidaat en toetst dus de eigen kennis van de student.

Geroosterde schriftelijke inzichttoets

Deze vindt plaats onder surveillance en er is door de beveiligde omgeving geen toegang tot AI.

Geroosterde schriftelijke kennistoets

Zie Geroosterde schriftelijke inzichttoets.

Geroosterde digitale inzichttoets

Geen risico met de randvoorwaarden dat er gebruik wordt gemaakt van een toetsapplicatie zoals ANS, fraudepreventiesoftware Schoolyear en surveillanten.

Geroosterde digitale kennistoets

Zie Geroosterde digitale inzichttoets.

Practicum live observatie

Zie Beroepshandeling (met live observatie)

Criteriumgericht interview (CGI)

De examiner stelt direct vragen aan de student. De student heeft hierbij geen toegang tot een laptop of het gebruik is zichtbaar voor de examiner. Het criteriumgericht interview is vaak aanvullend op een portfolio of product(en). Deze kunnen eventueel wel tot stand zijn gekomen door AI.

Practicum (met live observatie)

De risico's genoemd bij Practicum opname vallen weg door aanwezigheid van de docent.

A.2 MIDDELHOOG RISICO

Beroepshandeling (digitaal inleveren, opname)

Er zijn applicaties die tekst kunnen omzetten in film en in spraak. Deze applicaties worden steeds beter en toegankelijker voor een groter publiek, maar zijn dat op dit moment nog niet. Deze tools zullen in de toekomst wel een groter risico geven bij het digitaal inleveren van beroepshandelingen.

Presentatie

Het risico van gebruik van AI bij een presentatie is afhankelijk van wat je beoordeelt (inhoud of presentatievaardigheden). Bij het beoordelen van de inhoud, kan de inhoud gegenereerd zijn door AI. Bij het beoordelen van de presentatievaardigheden is de toetsing onder direct toezicht en is er geen risico. Vragen aan het einde van de presentatie kunnen verlagen het risico bij een beoordeling van de inhoud.

Practicum opname, zonder aanwezigheid docent

AI in te zetten voor de voorbereiding. AI inzetten voor vervalsen van de practicumopname is mogelijk, maar heeft nu nog een drempel. Zie Beroepshandeling.

Beroepsproduct (fysiek inleveren)

Bij een geschreven verslag heeft een beroepsproduct een hoog risico (zie A.3 Hoog Risico -> Beroepsproduct (digitaal inleveren)). Tastbare producten, zoals tekeningen of een model, hebben een lager risico. Bij het ontwerp of totstandkomingsproces kan wel AI gebruikt zijn.

Audiovisuele opdracht

Zie Beroepshandeling (digitaal inleveren, opname).

A.3 HOOG RISICO

Reflectieverslag

Een reflectieverslag komt meestal buiten het zicht van een examiner tot stand. Er ontbreekt hiermee aanvullende informatie om de authenticiteit te borgen.

Beoordelingsportfolio

Het risico van een beoordelingsportfolio is afhankelijk van de soort producten, LUK's, opdrachten, beoordelingscriteria en van het onderwijsproces (formatief handelen, zicht op student en de weg naar het portfolio toe, etc.). Veel portfolioproducten, zoals reflectieverslagen, maar ook beeldend materiaal laten zich makkelijk vervalsen door AI.

Beroepsproduct (digitaal inleveren)

Hoog risico als het een schriftelijk product betreft. Opgenomen handelingen (gesprek -geluid, video - presentatie) hebben een middelhoog risico. Deze opnames laten zich vervalsen met AI (bijvoorbeeld door Deep Fakes), maar deze tools zijn minder laagdrempelig toegankelijk en te gebruiken.

Schriftelijke opdracht met tijdslimiet

Schriftelijke opdrachten komen makkelijk tot stand met generatieve AI. Het risico wordt weggenomen als de opdracht wordt gemaakt in een klaslokaal met een toetsapplicatie zoals ANS en fraudepreventiesoftware Schoolyear en twee surveillanten.

Verantwoordingsverslag

Als een student een context meegeeft aan een prompt komt er snel een goed verslag uit een taalmodel rollen.

Schriftelijke opdracht

Zie Schriftelijke opdracht met tijdslimiet.

BIJLAGE B. TOETSACTIVITEITEN

Onderstaande activiteiten kunnen allemaal door AI worden uitgevoerd.

Welke activiteiten moet de student **voor de toets** ondernemen? Vink aan.

1. Tekst genereren en bewerken (verslag, rapport, plan, reflectie, advies, etc).

- (her)formuleren
- structureren
- samenvatten
- redigeren
- vertalen

2. Audio-visuele content creëren en opmaak verzorgen. Bijvoorbeeld het creëren van:

- audio
- video
- animatie
- poster
- flyer
- presentatie
- afbeelding
- infographic

3. Bijdragen aan (onderzoeks)opdrachten, zoals:

- brainstormen
- argumenteren en verantwoorden
- literatuur zoeken
- bronnen zoeken
- analyseren
- conclusies trekken
- reflecteren
- interviews genereren (fictieve gesprekken)

BIJLAGE C. LIJST MET MAATREGELLEN

Welke maatregelen in de onderwijs- en/of toetsactiviteiten neem jij om de toets AI-bestendig te maken?

Aanpassen leeruitkomsten

- De leeruitkomsten actualiseren, passend bij werkveld (visie op beroep) en visie op onderwijs.
 - ➔ Leg contact met de curriculumcommissie of dit mogelijk is.
- Het aanpassen van leeruitkomsten is een lange termijn maatregel in verband met einkwalificaties, curriculum, OS/OER, etc. waarbij ook het werkveld of de beroepenveldcommissie betrokken moet worden.

Maatregelen toetsing

- De toetsvorm aanpassen (een vorm van real time assessment), passend bij de leeruitkomsten en leeractiviteiten van de cursus en bij het toetsbeleid van de opleiding [hulpmiddel: Powerpoint AI in je toetsontwerp - Haagse Hogeschool]
- Beoordelingscriteria actualiseren.
- De toetsopdracht specifieker of authentieker maken (inhoud specifiek voor een situatie, organisatie, context).
- De beoordelingsprocedure aanscherpen (meer examinatoren, op meer momenten formatieve feedback).
- De toetsafname aanpassen (bijv. onder gecontroleerde omstandigheden).

Voor de bovenstaande maatregelen geldt:

- ➔ Vraag advies aan de toetscommissie en toetsdeskundige.
- ➔ Leg bij aanpassing van de toets de toets opnieuw voor aan de toetscommissie.
- ➔ Bepaal (in overleg met Curriculumcommissie/Academiemanager/OSOER redacteur) of de aanpassingen wel of niet in het huidige studiejaar kunnen worden doorgevoerd (binnen de vastgestelde, geldende OS/OER). De toetsvorm is bijvoorbeeld vastgelegd in de OS/OER.

Maatregelen onderwijs:

- Het zicht op het leerproces van de student verbeteren (formatief handelen inzetten).
- Verantwoord gebruik van AI in het onderwijs behandelen en begeleiden.
- Een kritische houding en ethisch en integer handelen verwerken in je lessen.

OPEN UP
NEW **HAN_** UNIVERSITY
OF APPLIED SCIENCES
HORIZONS.