

Toetsingskader AI funderend onderwijs

versie 1.0, 1 april 2026



Inhoudsopgave

Samenvatting	5
Inleiding	7
Kansen en kaders van AI in het onderwijs	7
AI-Verordening	8
Voorbehoud en status van dit toetsingskader.....	8
Wat is dit toetsingskader niet?	9
Eigen afweging en ethiek blijft uitgangspunt.....	10
Opbouw toetsingskader	10
1. Valt de AI-toepassing onder de AI-Verordening?	11
1. Is het een machinegebaseerd systeem?	11
2. Is het systeem ontworpen om met verschillende niveaus van autonomie te werken?	12
3. Kan het systeem adaptief zijn na implementatie? (LET OP: optioneel criterium).....	12
4. Heeft het systeem expliciete of impliciete doelstellingen?.....	13
5. Voert het systeem inferentie uit?	14
6. Genereert het systeem relevante outputs?	17
7. Kan het systeem fysieke of virtuele omgevingen beïnvloeden?	18
2. Is sprake van verboden AI-praktijken?.....	20
Verboden AI-praktijken	21
Toelichting emotieherkenning in het onderwijs	22
3. Is er sprake van een hoog-risico-AI-systeem?	24
A. Toegang tot onderwijsinstellingen.....	25
B. Evaluatie van leerresultaten	26
C. Bepalen van onderwijsniveau	28
D. Monitoring en detectie van ongeoorloofd gedrag tijdens toetsen	30
Overige voor onderwijs relevante hoog-risico-AI-systemen.....	31
Biometrie en Biometrische Identificatie	32

Toegang tot Essentiële Diensten	32
Werkgelegenheid en Werknemersbeheer.....	32
Rechtshandhaving en Individuele Risicobeoordeling	33
Uitzonderingen hoog-risico-AI	33
Profilering	35
Verschuiving van aanbiedersverantwoordelijkheid.....	35
4. Maatregelen voor hoog-risico-AI-systemen.....	38
Eisen voor hoog-risico-AI-systemen.....	38
Verplichtingen voor de aanbieder/leverancier	40
Verplichtingen voor de gebruiksverantwoordelijke/school.....	42
Fundamental Rights Impact Assessment (FRIA).....	45
Transparantieplichtingen	45
Bijlage 1: Ethische leidraad voor gebruik AI en data	47
Mensenrechten en ethiek	47
Ethische richtlijnen voor leveranciers en scholen.....	48
Bijlage 2: Checklist algoritmes	50
Bijlage 3: FAQ en aanbevelingen	53
Bronvermelding en verantwoording	56

Voor dit document hanteren wij classificatie: TLP: CLEAR

SIVON maakt gebruik van de volgende classificaties en coderingen:

Classificatie	Toelichting
TLP: RED	Strikte toegangscontrole. Opslag op beveiligde servers met encryptie. Distributie alleen via beveiligde kanalen.
TLP: AMBER	Dit document mag gedeeld worden binnen SIVON en (externe) belanghebbenden. Ontvanger mag niet verder delen.
TLP: GREEN	Geen restricties op toegang of distributie. Mag vrijelijk gedeeld worden binnen en buiten SIVON, maar niet via publieke kanalen.
TLP: CLEAR	Geen restricties op toegang of distributie. Mag vrijelijk gedeeld worden, ook via publieke kanalen binnen en buiten SIVON.

COLOFON - TOETSINGSKADER AI FUNDEREND ONDERWIJS

Auteurs: Job Vos (SIVON), Loes van Zijdam en Wietse van Bruggen (Kennisnet), Arjan Geurts (SIVON)

Met dank aan: MEWW, VDOD, EDU-V, SURF, PO-Raad/VO-raad, NOLAI en NPULS

Versie: 1.0 (1 april 2026)

Status: *voorlopig*

Hoewel aan dit document de uiterste zorg is besteed, aanvaarden SIVON en de auteurs geen aansprakelijkheid voor eventuele fouten, onvolkomenheden en zijn evenmin aansprakelijkheid voor schade van welke aard ook, ontstaan uit het (onjuiste) gebruik van dit document.



Het Toetsingskader AI is opgesteld door SIVON en verschijnt onder de licentie Creative Commons Naamsvermelding 4.0 Nederland.

Heb je vragen over dit document? Kijk voor meer informatie op

<http://www.sivon.nl/toetsingskader-ai> of stuur een mail naar ibp@sivon.nl

Dit Toetsingskader AI is opgesteld door SIVON en mogelijk gemaakt door een bijdrage van het ministerie van OCW.

Samenvatting

Voorbehoud: dit toetsingskader is gebaseerd op informatie tot en met 9 maart 2026. De nog te publiceren uitleg van de Europese Commissie over AI-systemen met een hoog risico en de voorgestelde maatregelen uit de Digitale Omnibus worden later dit jaar verwerkt in een volgende versie. De status van dit document is: *voorlopig*.

Artificiële intelligentie (AI) biedt scholen kansen voor onderwijsinnovatie, ondersteuning van leerprocessen en vermindering van administratieve lasten. Tegelijk brengt AI risico's mee voor privacy, gegevensbescherming, transparantie, gelijke behandeling en de onderwijskwaliteit.

Het toetsingskader richt zich in eerste plaats op leveranciers. Zij kunnen hiermee, op basis van de AI-Verordening, nagaan of hun ict-systeem met AI onder de nieuwe regels valt, welke risico's dat meebrengt en welke maatregelen nodig zijn om aan de wettelijke eisen te voldoen en risico's voor leerlingen en medewerkers te beperken. Daaruit volgt dat scholen met dit toetsingskader kunnen beoordelen of de leverancier de vereiste maatregelen heeft genomen én welke maatregelen de school zelf nog moet treffen.

Dit Toetsingskader AI voor het funderend onderwijs vertaalt de AI-Verordening naar de onderwijspraktijk. Het toetsingskader is gebaseerd op de AI-Verordening (EU 2024/1689), gericht op leveranciers in de educatieve sector en schoolbesturen in Nederland, schept geen nieuwe verplichtingen en biedt partijen een eenduidige uitleg van kernbegrippen uit de AI-Verordening.

Het doel van het toetsingskader is:

1. identificeren of een AI-systeem onder de AI-Verordening valt,
2. bepalen of een AI-systeem verboden is, in de categorie hoog risico valt, of onder een uitzondering valt.
3. op basis van de uitkomst de verplichtingen voor leveranciers en scholen in kaart brengen.

De te doorlopen vragen en stappen zijn hierna opgenomen. Hierbij geldt de kanttekening dat de uitleg en duiding in het betreffende hoofdstuk en paragraaf zijn opgenomen. Het toetsingskader biedt richting voor het funderend onderwijs, maar de individuele beoordeling van een AI-systeem blijft steeds leidend.

1. Valt het AI-systeem onder de AI-Verordening?

Hiervoor gelden volgens de definitie de volgende zeven criteria:

1. Is het een machine-, logica- en/of kennisgebaseerd systeem?
2. Is het systeem ontworpen om met verschillende niveaus van autonomie te werken?
3. Kan het systeem adaptief zijn na implementatie? (LET OP: optioneel criterium)
4. Heeft het systeem expliciete of impliciete doelstellingen?
5. Voert het systeem inferentie uit?
6. Genereert het systeem relevante outputs?
7. Kan het systeem fysieke of virtuele omgevingen beïnvloeden?

2. Is sprake van een verboden AI-praktijken?

3. Is sprake van een hoog-risicotoepassing?

- AI-systeem voor toelating tot of toewijzing van (beroeps)onderwijs
- AI-systeem voor het evalueren van leerresultaten
- AI-systeem voor het beoordelen van het (beroeps)onderwijsniveau
- AI-systeem voor het monitoren van deelnemers tijdens examens

4. Is sprake van een uitzondering op de hoog-risicocategorie?

- a. Een beperkte procedurele taak uitvoeren
- b. Het verbeteren van een eerder voltooide menselijke activiteit
- c. Het opsporen van besluitvormingspatronen of afwijkingen, zonder de menselijke beoordeling te vervangen of te beïnvloeden
- d. Een voorbereidende taak uitvoeren voor een beoordeling die relevant is voor hoog-risicogebruiksgevallen

Let op: indien er sprake is van profilering dan gelden deze uitzonderingen niet.

Inleiding

Kansen en kaders van AI in het onderwijs

Artificiële intelligentie (AI) biedt grote kansen voor het onderwijs: van gepersonaliseerd leren en efficiëntere administratie tot innovatieve leermiddelen. Tegelijk brengt AI risico's mee, zoals privacyschendingen, discriminatie door algoritmische vooroordelen en ondoorzichtige besluitvorming die de autonomie van leerlingen en docenten kan aantasten. Om kansen te benutten én risico's te beheersen, is helder inzicht nodig in de verplichtingen en verantwoordelijkheden rond de inzet van AI. De Europese AI-Verordening (AI Act) vormt daarbij het centrale wettelijke kader voor de beoordeling van AI-toepassingen, de risicoclassificatie daarvan en de daaruit voortvloeiende verplichtingen voor leveranciers en scholen.

Het is essentieel dat leveranciers en scholen de AI-Verordening op dezelfde wijze uitleggen en toepassen. Dit Toetsingskader AI voor het funderend onderwijs helpt leveranciers en scholen te beoordelen of AI-systemen voldoen aan wettelijke eisen, zoals de AVG en de AI-Verordening, en of zij ethisch verantwoord en veilig kunnen worden ingezet. Zo borgt de inzet van AI niet alleen innovatie, maar ook de bescherming van grondrechten, gelijkheid en transparantie in het onderwijs. Het toetsingskader helpt vast te stellen of een toepassing van artificiële intelligentie in een ict-systeem of digitaal onderwijsmiddel, zoals digitaal leermateriaal of een leerlingadministratiesysteem, onder de AI-Verordening valt en welke risicoclassificatie daarop van toepassing is.

Het toetsingskader richt zich in eerste plaats op leveranciers. Zij kunnen hiermee, op basis van de AI-Verordening, nagaan of hun ict-systeem met AI onder de nieuwe regels valt, welke risico's dat meebrengt en welke maatregelen nodig zijn om aan de wettelijke eisen te voldoen en risico's voor leerlingen en medewerkers te beperken. Daaruit volgt dat scholen met dit toetsingskader kunnen beoordelen of de leverancier de vereiste maatregelen heeft genomen én welke maatregelen de school zelf nog moet treffen.

Dit Toetsingskader AI voor het funderend onderwijs vertaalt de AI-Verordening naar de Nederlandse onderwijspraktijk. Het kader schept geen nieuwe verplichtingen en biedt partijen een eenduidige uitleg van kernbegrippen uit de AI-Verordening.

AI-Verordening

De AI-Verordening kent een risicogebaseerde benadering voor ict-systemen en software waarin AI wordt toegepast. Afhankelijk van de risicoclassificatie gelden verschillende verplichtingen voor leveranciers en gebruikers. De verordening beoogt innovatie en het gebruik van AI te stimuleren, terwijl tegelijk gezondheid, veiligheid en fundamentele rechten worden beschermd. Of een toepassing onder de AI-Verordening valt, wordt in dit toetsingskader beoordeeld aan de hand van stappen die gebaseerd zijn op de richtsnoeren van de Europese Commissie over de definitie van een AI-systeem.

De AI-Verordening is op 1 augustus 2024 in werking getreden en wordt gefaseerd van kracht: Voor AI-systemen in ontwikkeling geldt dat de verordening volledig in werking treedt op 2 augustus 2026, wat betekent dat deze systemen vanaf die datum aan de voorschriften moeten voldoen. Voor AI-systemen in gebruik die vóór 2 augustus 2026 zijn geïntroduceerd, geldt dat de verordening alleen van toepassing is als ze significante wijzigingen ondergaan, maar voor AI-systemen met een hoog risico, met name die bedoeld voor overheidsgebruik, moeten aanbieders uiterlijk op 2 augustus 2030 aan alle vereisten voldoen, ongeacht ontwerpwijzigingen.

In dit toetsingskader worden de termen aanbieder en leveranciers afwisselend gebruikt: het gaat hierbij om de partij die een AI-systeem ontwikkelt (laat ontwikkelen) en dat systeem in de handel brengt of in gebruik stelt onder eigen naam. Met gebruiksverantwoordelijke worden de schoolbesturen bedoeld, de partij die een AI-systeem onder eigen verantwoordelijkheid gebruikt. Daarnaast kan er een importeur zijn die een AI-systeem van buiten naar binnen de EU importeert. De distributeur is de partij die in de toeleverketen die ene AI-systeem in de EU op de markt aanbiedt.

Voorbehoud en status van dit toetsingskader

Dit toetsingskader is gebaseerd op de stand van zaken *tot en met 9 maart 2026*. Nieuwe uitleg van de Europese Commissie over hoog risico AI¹ en eventuele wijzigingen (zoals uit de Digitale Omnibus²) worden in latere versies verwerkt. Het kader is daarmee nadrukkelijk dynamisch, heeft een voorlopig karakter en wordt, na nieuwe informatie is gepubliceerd, in 2026 door SIVON bijgewerkt.

¹ <https://ai-act-service-desk.ec.europa.eu/en/faq> : The Commission is currently preparing guidelines on the classification of high-risk AI systems, which should be published in February 2026.

² <https://digital-strategy.ec.europa.eu/en/library/digital-omnibus-ai-regulation-proposal>

Het toetsingskader biedt richting voor het funderend onderwijs, maar de individuele beoordeling van een AI-systeem blijft steeds leidend. Het wijst daarom niet categorisch aan welke systemen wel of niet onder de AI-Verordening vallen of welke AI-systemen kwalificeren als hoog risico.

Wat is dit toetsingskader niet?

Dit toetsingskader richt zich op educatieve AI³ in ict-systemen en software voor scholen. Algemene of generatieve AI-modellen⁴ vallen buiten de primaire focus, tenzij zij in de praktijk worden ingezet als educatieve AI. Het kader gaat over leren mét of ondersteund door AI, niet over leren óver AI (AI-geletterdheid). Ook beoordeelt het niet of de inzet van AI past binnen de visie of het beleid van de school of pedagogisch geschikt is. Daarvoor kan onder meer de [Handreiking AI in het onderwijs](#) van Kennisnet worden gebruikt.

Niet elk algoritme valt onder de definitie van een AI-systeem in de zin van de AI-Verordening. Dat een leverancier zijn product als "AI" aanduidt, is dus geen garantie dat de verordening van toepassing is. Dit toetsingskader richt zich primair op AI-systemen die leveranciers in hun producten hebben ingebouwd, niet op losse AI-tools of generatieve taalmodellen zoals Copilot, Claude of ChatGPT (General-Purpose AI, GPAI). Voor die laatste gelden overigens wel aparte verplichtingen in de AI-Verordening.

Een aandachtspunt vormt het gebruik van generatieve AI voor toepassingen die mogelijk onder de hoog-risicocategorieën vallen. In dat geval bestaat de kans dat een school wordt aangemerkt als aanbieder in de zin van de verordening, met de bijbehorende verplichtingen.

Valt een toepassing buiten de AI-Verordening, dan blijven andere kaders, zoals de AVG, cyberveiligheidsregels en onderwijswetgeving onverminderd van toepassing en kunnen algoritmische risico's voor de rechten en vrijheden van betrokkenen nog steeds aan de orde zijn. Soms is een vorm van geautomatiseerde besluitvorming onder omstandigheden geen hoog risico, maar daarop kan wel het verbod van geautomatiseerde besluitvorming (art 22 AVG) van toepassing zijn.

³ Educational AI, in: *Updated Guidelines on the ethical use of artificial intelligence and data in teaching and learning for educators* (Europese Commissie November 2025, updated March 2026)

⁴ GPAI: general purpose AI

Eigen afweging en ethiek blijft uitgangspunt

Scholen en leveranciers moeten blijven beoordelen of de inzet van AI rechtvaardig, transparant, uitlegbaar en onder menselijke controle plaatsvindt, zonder bias of discriminatie. AI in het onderwijs roept niet alleen juridische en technische, maar ook ethische vragen op. AI moet rechtvaardig, transparant en inclusief worden ingezet, met respect voor de fundamentele rechten en waardigheid van leerlingen en medewerkers. In het bijzonder vraagt dit aandacht voor bias en discriminatie, transparantie en uitlegbaarheid en menselijke controle. De AI-Verordening gaat er van uit dat menselijk toezicht altijd mogelijk blijft.

In bijlage 1 zijn de ethische richtsnoeren voor gebruik van AI in het onderwijs opgenomen, deze geven richting ook voor die gevallen dat er geen sprake is van hoog risico of een AI-toepassing niet onder de AI-Verordening valt.

Opbouw toetsingskader

Het toetsingskader volgt de stappen die nodig zijn om te beoordelen of een AI-systeem onder de AI-Verordening valt en welke verplichtingen daaruit voortvloeien.

Achtereenvolgens komen aan bod: de AI-definitie, verboden AI-toepassingen, hoog-risico-AI en de uitzonderingen daarop, en de verplichtingen voor *hoog-risico-AI-systemen*, leveranciers en scholen.

De bijlagen bieden verdieping op een aantal aangrenzende thema's. Bijlage 1 werkt de ethische richtsnoeren voor AI in het onderwijs verder uit. Bijlage 2 bevat een check voor algoritmen, ook bruikbaar voor systemen die buiten de AI-Verordening vallen. Als laatste zijn in bijlage 3 veelgestelde vragen opgenomen. Tot slot zijn de bronnen en verwijzingen opgenomen in een apart hoofdstuk.

1. Valt de AI-toepassing onder de AI-Verordening?

Om te bepalen of een AI-toepassing of ict-systeem met AI, onder de AI-Verordening valt, moet worden voldaan aan de definitie van artikel 3 lid 1 van de AI-Verordening.

“AI-systeem”: een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na het inzetten ervan aanpassingsvermogen kan vertonen, en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren zoals voorspellingen, inhoud, aanbevelingen of beslissingen die van invloed kunnen zijn op fysieke of virtuele omgevingen.

Deze definitie is op te delen in zeven aparte onderdelen waarop kan worden getoetst. Als wordt voldaan aan deze zeven criteria, valt de AI-toepassing onder de AI-Verordening en is er sprake van een ‘AI-systeem’. Hierbij geldt dat er niet op alle momenten tegelijk aan alle criteria moet worden voldaan: er is de fase van het ontwikkelen van het AI-systeem en de fase van het inzetten van het AI-systeem. Ook moeten de elementen van de definitie niet mechanisch⁵ worden toegepast maar op basis van de specifieke architectuur en functionaliteit van ieder specifiek systeem.

Het gaat om de volgende zeven criteria:

1. Is het een machinegebaseerd systeem?

Hulpvragen⁶:

- Bevat het systeem hardwarecomponenten (zoals processoren, geheugen, opslag, netwerkeenheden, input/output interfaces) die de rekeninfrastructuur leveren?
- Bevat het systeem softwarecomponenten (zoals computercode, instructies, programma's, besturingssystemen, applicaties) die bepalen hoe de hardware data verwerkt en taken uitvoert?
- Is het systeem computationeel gedreven en afhankelijk van machinebewerkingen?
- Wordt het systeem uitgevoerd op fysieke of cloud-gebaseerde machines?

⁵ Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act), overweging 6.

⁶ Het is bij alle 7 criteria niet noodzakelijk dat alle hulpvragen positief beantwoord zijn om onder de definitie te vallen.

Uitzondering: systemen die volledig handmatig werken zonder computationele verwerking, of systemen die zijn ontworpen om uitsluitend met volledige handmatige menselijke betrokkenheid en interventie te werken.

Uitleg: een rekenmachine is machinegebaseerd, maar geen AI. Een rekenmachine die zelf leert welke sommen een leerling moeilijk vindt en daarop aangepaste oefeningen suggereert, wel. Het verschil is dat de laatste zelfstandig redeneert op basis van data.

2. Is het systeem ontworpen om met verschillende niveaus van autonomie te werken?

Hulpvragen:

- Is het systeem technisch zo ontworpen om met autonomie te werken?
- Kan het systeem bijvoorbeeld zelfstandig output genereren?
- Kan het systeem functioneren met enige mate van onafhankelijkheid dus zonder directe menselijke tussenkomst?

NB “enige mate van autonomie” kan ook eenvoudige automatisering omvatten. Het gaat bij dit criterium om technisch ontwerp en niet om daadwerkelijk gebruik.

Uitzondering: systemen die ontworpen zijn om alleen te functioneren met volledige menselijke sturing (bijv. een docent die elke stap handmatig invoert en bevestigt).

Uitleg: een auto met adaptive cruisecontrol (die zelf de snelheid aanpast op basis van de afstand tot de voorligger) is ontworpen voor autonomie (AI), ook als de bestuurder het nooit gebruikt. Een auto met gewone cruisecontrol is niet ontworpen voor autonomie (geen AI), ook niet als de bestuurder de auto de hele weg laat rijden.

3. Kan het systeem adaptief zijn na implementatie? (LET OP: optioneel criterium)

Hulpvragen:

- Is het systeem ontworpen om zich aan te passen aan nieuwe input of omstandigheden tijdens gebruik, beschikt het over zelflerende capaciteiten?
- Kan het systeem verschillende resultaten produceren voor dezelfde inputs na adaptatie (bijvoorbeeld afhankelijk van context of gebruikersgedrag)?

- Wijzigt het systeem tijdens gebruik zijn gedrag op basis van nieuwe input (bijv. leerlingprestaties, gebruikersinteracties) zonder dat dit gedrag vooraf is geprogrammeerd?

Uitzondering: dit criterium is niet verplicht; een systeem kan onder de AI-Verordening vallen zonder adaptiviteit.

Uitleg adaptiviteit: een thermostaat die zelf leert wanneer jij thuis bent en de temperatuur daarop aanpast, is adaptief. Een thermostaat met vaste instellingen is niet adaptief (geen AI). Het verschil is dat de eerste zich aanpast aan nieuwe input (wanneer wel of niet thuis).

4. Heeft het systeem expliciete of impliciete doelstellingen?

Hulpvragen:

- Zijn er expliciete doelstellingen die rechtstreeks in het systeem zijn vastgelegd, zoals het optimaliseren van een kostenfunctie, kans, score of beloning?
- Zijn er impliciete doelstellingen die niet letterlijk zijn geprogrammeerd, maar kunnen worden afgeleid uit het gedrag van het systeem, de trainingsdata of de interactie met de omgeving?
- Zijn er impliciete doelen die voortkomen uit het gedrag van het systeem of de interactie met zijn omgeving (bijv. inzicht over klasniveaus door gebruik)?

N.B. De doelstellingen van het AI-systeem moeten worden onderscheiden van het beoogde gebruik (intended purpose) van de aanbieder. Doelstellingen zijn intern aan het systeem en zien toe op de taak en het resultaat dat het systeem probeert te bereiken. Beoogd gebruik is extern en ziet toe op het gebruik waarvoor de aanbieder het systeem bedoelt, inclusief de gebruikscontext, voorwaarden en instructies voor gebruik.

Uitzondering: n.v.t.

Uitleg: een navigatie-app heeft expliciet doelen zoals het minimaliseren van de reistijd en afstand, het optimaliseren van de routekosten. Het impliciete doel kan zijn het bijhouden van routevoorkeuren om daarmee routes aan te passen aan individuele voorkeuren. Het beoogd doel is bijvoorbeeld het ondersteunen van gebruikers bij het plannen en volgen van hun route naar hun bestemming in wegverkeer.

5. Voert het systeem inferentie uit?

Inferentie verwijst naar het vermogen van een AI-systeem om output (zoals voorspellingen, aanbevelingen of beslissingen) af te leiden uit input. Dit kan zowel plaatsvinden tijdens de ontwikkelfase (bijvoorbeeld het afleiden van modellen of regels uit data) als tijdens de gebruiksfase (het toepassen van die modellen of regels om output te genereren). Een systeem voert inferentie uit wanneer het, op basis van data of gecodeerde kennis, conclusies trekt, voorspellingen doet of nieuwe informatie genereert, en daarmee verder gaat dan het louter uitvoeren van vooraf vastgelegde instructies.

Hulpvragen:

- Gebruikt het systeem technieken die inferentie mogelijk maken?
 - Machine Learningtechnieken (tijdens ontwikkelfase):
 - Supervised learning (bijv. nakijksoftware die leert om fouten te herkennen).
 - Unsupervised learning (bijv. clustering van toetsresultaten om leerbehoeften te identificeren).
 - Self-supervised learning (bijv. taalmodellen die patronen in tekst leren herkennen).
 - Reinforcement learning (bijv. adaptieve leerpaden die zich aanpassen aan leerlinggedrag).
 - Deep learning (bijv. spraakherkenning voor taalonderwijs).
 - Logica- en kennisgebaseerde technieken (tijdens ontwikkelfase):
 - Inferentie-engines (bijv. expertsystemen die schooladvies genereren op basis van regels).
 - Symbolische redenering: het systeem gebruikt logische regels (bijv. 'als X, dan Y') om conclusies te trekken (bijv. wiskundige software die stapsgewijs redeneren stimuleert).
 - Kennisrepresentatie: het systeem slaat kennis op in een gestructureerde vorm (bijv. semantische netwerken) om redeneringen uit te voeren, zoals een digitale leeromgeving die concepten met elkaar verbindt.
- Voert het systeem inferentie uit tijdens gebruik?
 - Past het systeem zijn gedrag aan op basis van nieuwe input (bijv. een adaptief leersysteem dat opgaven kiest op basis van leerlingprestaties)?
 - Genereert het systeem verschillende outputs voor dezelfde input, afhankelijk van context (bijv. een chatbot die antwoorden aanpast aan de vraagsteller)?

- Leidt het systeem op basis van input nieuwe (tussen)conclusies af voordat het tot een uitkomst komt (bijv. een nakijktool die foutpatronen analyseert en daaruit afleidt welke feedback nodig is)?
- Wijzigt het systeem realtime zijn gedrag op basis van nieuwe input tijdens gebruik, zonder dat dit vooraf is geprogrammeerd?
 - a. Is het een logica- of kennis-gebaseerd systeem dat redeneert op basis van gecodeerde regels, kennis, feiten, etc.?

Uitzonderingen: de volgende systemen vallen in beginsel buiten de AI-definitie, omdat zij doorgaans geen inferentie uitvoeren in de zin van de AI-Verordening, of omdat zij wel een beperkte inferentiecapaciteit hebben maar basic data processing niet overstijgen. Een zorgvuldige beoordeling van de concrete werking van het systeem blijft nodig.

- Systemen voor basisgegevensverwerking (bijv. standaard spreadsheets, query-systemen die gegevens ophalen of berekenen zonder interpretatie).

Hulpvragen:

- Voert het systeem uitsluitend vooraf gedefinieerde bewerkingen uit op data (sorteren, filteren, optellen)?
- Genereert het systeem uitsluitend output op basis van regels die volledig door een mens zijn vastgelegd, zonder enige patroonherkenning of redenering?
- Systemen waarmee wiskundige optimalisatie wordt verbeterd of traditionele optimalisatiemethoden (zoals lineaire of logistische regressie) worden versneld of benaderd, voor zover zij de elementaire verwerking van data niet overstijgen.

Hulpvragen:

- Is het systeem gericht op het verbeteren van de computationele prestaties van een bestaand optimalisatieproces?
- Blijft de werking functioneel vergelijkbaar met traditionele, gevestigde optimalisatiemethoden?
- Overstijgt het systeem de elementaire dataverwerking, bijvoorbeeld doordat het zelfstandig betekenisvolle patronen analyseert of zijn beslismodel intelligent aanpast?
- Worden eventuele zelfaanpassingen alleen gebruikt om de computationele efficiëntie te verbeteren, en niet om het beslismodel inhoudelijk intelligenter te maken?
- Systemen die gebaseerd zijn op uitsluitend door natuurlijke personen gedefinieerde regels om automatisch operaties uit te voeren.

Hulpvragen:

- Zijn alle regels, vertakkingen en uitkomsten volledig vooraf door mensen vastgelegd?
- Kan het systeem geen conclusies trekken buiten de vooraf gedefinieerde regelset?
- Bevat het systeem geen mechanisme om op basis van data nieuwe regels, gewichten of relaties af te leiden?
- Past het systeem geen logica- of kennisgebaseerde inferentie toe om nieuwe situaties te beoordelen? Eenvoudige voorspellingssystemen (bijv. het berekenen van een gemiddelde of trend op basis van historische waarden, zonder gebruik van AI-technieken).

Hulpvragen:

- Berust de voorspelling op een vaste, door mensen gedefinieerde statistische basisregel (bijv. voortschrijdend gemiddelde)?
- Herkent het systeem geen complexe patronen in data en past het zijn voorspellingsmethode niet aan op basis van nieuwe input?
- Klassieke heuristische systemen (bijv. systemen die met vooraf gedefinieerde heuristische regels of algoritmen benaderende oplossingen zoeken, zoals een schaakprogramma met minimax en heuristische evaluatiefuncties)

Hulpvragen:

- Past het systeem vooraf gedefinieerde heuristische regels of algoritmen toe om efficiënt een benaderende oplossing te vinden?
- Berust de werking op klassieke programmeertechnieken, zoals rule-based benaderingen, trial-and-error-strategieën of heuristische evaluatiefuncties, in plaats van op datagedreven leren?
- Leert het systeem niet uit trainingsdata of uit input-outputrelaties op een wijze die het model inhoudelijk aanpast?
- Past het systeem vooraf vastgelegde oplossingsregels toe zonder zelf nieuwe modellen, gewichten of relaties uit data af te leiden?

Let op:

- Niet elke machine learning toepassing is AI: traditionele optimalisatiemodellen (bijv. lineaire regressie voor eenvoudige voorspellingen) vallen buiten de definitie als ze geen geavanceerde inferentie uitvoeren.

- Grensgevallen vereisen zorgvuldige beoordeling (bijv. klassieke heuristiek vs. symbolische AI).
- Inferentie tijdens gebruik verhoogt het risico: systemen die realtime aanpassingen maken, vallen vaak onder strengere regels (bijv. hoog-risico-AI).

Uitleg: banksoftware die op basis van input als leeftijd, opleiding, woonplaats en zoekgeschiedenis van een gebruiker, een aanbeveling genereert voor een financiële dienst (bijvoorbeeld een specifiek type lening), zonder dat die uitkomsten specifiek voorgeprogrammeerd zijn. Op basis van input genereert het systeem output op basis van patroonherkenning of statistische analyse.

6. Genereert het systeem relevante outputs?

Hulpvragen:

- Genereert het systeem voorspellingen — dat wil zeggen: schattingen van een onbekende waarde op basis van bekende inputwaarden (bijv. een inschatting van de kans op schooluitval, of een voorspeld eindcijfer)?
- Genereert het systeem inhoud — dat wil zeggen: nieuw gegenereerd materiaal zoals tekst, afbeeldingen, audio of video? Dit type output komt in het bijzonder voor bij generatieve AI-systemen, zoals systemen die toetsvragen opstellen, leerlingfeedback formuleren of lesmateriaal samenstellen.
- Doet het systeem aanbevelingen — dat wil zeggen: suggesties voor een handeling of keuze (bijv. een gepersonaliseerd leerpad, een advies voor extra ondersteuning, een school- of profielkeuze)? Het maakt daarbij niet uit of de aanbeveling daadwerkelijk wordt opgevolgd: ook een aanbeveling die niet tot actie leidt, is output in de zin van de AI-Verordening.
- Neemt het systeem beslissingen — dat wil zeggen: conclusies of keuzes die door het systeem worden getrokken waarbij processen zijn geautomatiseerd die traditioneel door mensen worden verricht, en waarbij een bepaald resultaat zonder menselijke tussenkomst wordt geproduceerd (bijv. geautomatiseerde toelating, automatische klasindeling)?

N.B. AI-systemen onderscheiden zich van niet-AI-systemen niet zozeer doordat zij output genereren, maar door de wijze waarop die output tot stand komt. De AI-Verordening noemt als mogelijke outputs: voorspellingen, content, aanbevelingen en beslissingen. Het onderscheid met traditionele software ligt erin dat AI-systemen op basis van input kunnen infereren hoe zulke output moet

worden gegenereerd, terwijl niet-AI-systemen doorgaans werken met statische gegevens en vooraf vastgestelde regels.

Uitzondering: n.v.t.

Uitleg: een systeem dat leerresultaten in een tabel weergeeft op basis van ingevoerde cijfers, genereert output maar voert geen inferentie uit: geen AI. Een systeem dat op basis van diezelfde resultaten een patroon herkent en een gepersonaliseerd leerp pad voorstelt, genereert output op basis van inferentie: wel AI. Het verschil zit niet in wat er met de output gebeurt, maar in hoe die output tot stand is gekomen.

7. Kan het systeem fysieke of virtuele omgevingen beïnvloeden?

Hulpvragen:

- Produceert het systeem voorspellingen die fysieke of virtuele omgevingen kunnen beïnvloeden, bijvoorbeeld een voorspelling over uitvalrisico of verwachte leerprestaties?
- Doet het systeem aanbevelingen die keuzes of handelen van gebruikers kunnen beïnvloeden, bijvoorbeeld adviezen voor klassenindeling of extra leerondersteuning?
- Neemt het systeem beslissingen die fysieke processen aansturen of wijzigen, bijvoorbeeld via een robot of slimme thermostaat?
- Neemt het systeem beslissingen die digitale processen, datastromen of softwaregedrag wijzigen, bijvoorbeeld door leerpaden aan te passen, activiteiten te blokkeren of netwerkverkeer te prioriteren?
- Genereert het systeem inhoud, zoals tekst, beeld of audio, die fysieke of virtuele omgevingen kan beïnvloeden, bijvoorbeeld een lesplan, presentatie, afbeelding of gesproken uitleg?

NB met “beïnvloeden” wordt bedoeld dat de output van het systeem effect kan hebben op fysieke objecten, processen of op virtuele omgevingen zoals digitale werkruimten, data- en softwareprocessen. Het is niet vereist dat elke output automatisch en zonder menselijke tussenkomst wordt uitgevoerd.

Uitzondering: “niet-autonome systemen die zich beperken tot het tonen van informatie zonder verdere actie?”

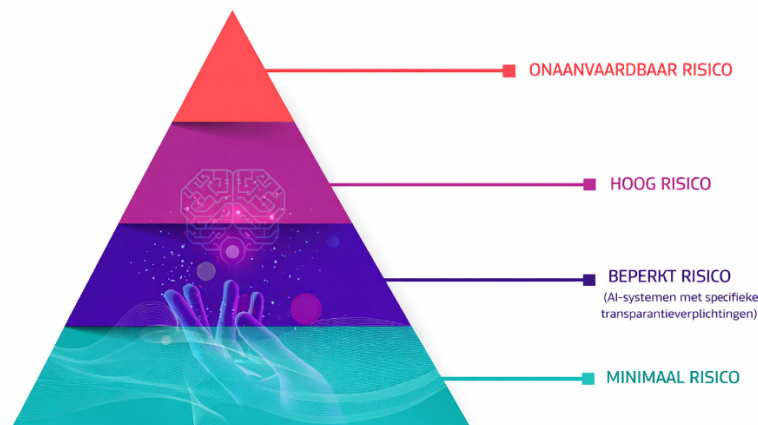
Uitleg: een systeem dat op basis van leerlingdata een uitvalrisico als score op een dashboard toont, kan een AI-systeem zijn, omdat het een voorspelling genereert op basis van inferentie. Een systeem dat alleen ingevoerde cijfers optelt en een gemiddelde toont, is in beginsel geen AI-systeem, omdat het slechts een vooraf vastgelegde rekenregel uitvoert. Beide tonen informatie, maar zij verschillen in de wijze waarop die output tot stand komt.

Voldoet een systeem aan alle zeven criteria (waarbij stap 3 facultatief is), dan kwalificeert het als een AI-systeem in de zin van artikel 3, lid 1, AI-Verordening en valt het binnen de reikwijdte van de AI-Verordening. Daarna moet afzonderlijk worden beoordeeld welke bepalingen van toepassing zijn. Slechts een deel van de AI-systemen valt onder verboden praktijken, hoog-risicoverplichtingen of transparantieplichtingen; de grote meerderheid niet. Wel kunnen algemene bepalingen, zoals de verplichting tot AI-geletterdheid voor aanbieders en gebruiksverantwoordelijken relevant blijven.

Een AI-systeem kan een afgebakend onderdeel zijn van een groter ict-pakket, maar ook het geïntegreerde geheel. De kwalificatie en verplichtingen kleven aan het juridisch relevante AI-systeem, ongeacht of het fysiek geïntegreerd is of losgekoppeld functioneert. De relevante vraag bij de beoordeling van bijvoorbeeld een LAS is daarom niet "is dit hele pakket een AI-systeem en hoog risico?" maar "welke AI-component zit hierin, en wat is het beoogde gebruik daarvan en hoe kwalificeert dat onder de AI-Verordening?".

2. Is sprake van verboden AI-praktijken?

De AI-Verordening hanteert een risicogebaseerde aanpak: hoe groter het risico van een AI-systeem, hoe zwaarder de eisen. De verordening onderscheidt drie categorieën. AI-toepassingen die een onaanvaardbaar risico vormen voor gezondheid, veiligheid of grondrechten zijn verboden. Voor hoog-risico-AI-systemen gelden uitgebreide verplichtingen voor leveranciers en scholen.



7

De AI-Verordening van de Europese Unie hanteert een risicogebaseerde benadering waarbij AI-systemen worden ingedeeld in vier categorieën.

Bepaalde AI-praktijken brengen een onaanvaardbaar risico met zich mee voor mensen en de samenleving en zijn daarom verboden. Binnen de categorie verboden AI, zijn sommige AI-systemen per definitie verboden (ongeacht het gebruik). Bij andere AI-systemen is een specifieke toepassing (AI-praktijk) verboden omdat die onaanvaardbaar is, terwijl het onderliggende AI-systeem op zichzelf is toegestaan. Hierbij spelen het beoogde doel en de expliciete en impliciete doelstellingen dus een rol.

De AI-Verordening treedt gefaseerd in werking. Voor verboden AI-praktijken als bedoeld in artikel 5 geldt dat deze regels sinds 2 februari 2025 van toepassing zijn. Dat betekent dat dergelijke AI-praktijken vanaf die datum niet meer rechtmatig op de markt mogen worden gebracht, in gebruik mogen worden gesteld of mogen worden gebruikt.

Deze verboden gelden zowel voor aanbieders als gebruiksverantwoordelijken.

⁷ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

Verboden AI-praktijken

- a) Schadelijke manipulatie of misleiding: AI-systemen die subliminale technieken of doelbewust manipulatieve of misleidende technieken gebruiken, waardoor het vermogen van personen om een geïnformeerd besluit te nemen wezenlijk wordt aangetast en significante schade ontstaat of redelijkerwijs kan ontstaan.
- b) Schadelijke exploitatie van kwetsbaarheid: AI-systemen die kwetsbaarheden uitbuiten wegens leeftijd, handicap of een specifieke sociale of economische situatie, waardoor gedrag wezenlijk wordt verstoord en significante schade ontstaat of redelijkerwijs kan ontstaan.
- c) Sociale scoring: AI-systemen die personen of groepen over een bepaalde periode evalueren of classificeren op basis van sociaal gedrag of bekende, afgeleide of voorspelde persoonlijke of persoonlijkheidskenmerken, wanneer dat leidt tot ongunstige behandeling in een andere sociale context of tot ongerechtvaardigde of onevenredige ongunstige behandeling.
- d) Criminele risicovoorspelling uitsluitend op basis van profilering of persoonlijkheidskenmerken: AI-systemen die het risico voorspellen dat een natuurlijke persoon een strafbaar feit pleegt, voor zover dat uitsluitend is gebaseerd op profilering of op de beoordeling van persoonlijkheidskenmerken. Dit verbod geldt niet voor systemen die slechts menselijke beoordeling ondersteunen en daarbij steunen op objectieve en verifieerbare feiten die direct samenhangen met criminele activiteit.
- e) Ongerichte scraping voor gezichtsherkenningsdatabases: AI-systemen die door ongerichte scraping van internet of CCTV-beelden databases voor gezichtsherkenning creëren of uitbreiden.
- f) Emotie-afleiding op het werk of in het onderwijs: AI-systemen die emoties van natuurlijke personen afleiden in werkplekken of onderwijsinstellingen, behalve wanneer dat gebeurt voor medische of veiligheidsdoeleinden.
- g) Biometrische categorisering om beschermde kenmerken af te leiden: AI-systemen die individuen op basis van biometrische gegevens categoriseren om daaruit ras, politieke opvattingen, vakbondslidmaatschap, religieuze of filosofische overtuigingen, seksleven of seksuele gerichtheid af te leiden, behoudens de beperkte uitzonderingen uit de verordening.
- h) Real-time biometrische identificatie (RBI) op afstand in publiek toegankelijke ruimten voor rechtshandhaving: AI-systemen voor real-time biometrische identificatie op afstand in publiek toegankelijke ruimten ten behoeve van rechtshandhaving, behoudens de zeer beperkte uitzonderingen en strikte voorwaarden van de verordening.

Toelichting emotieherkenning in het onderwijs

De definitie van systemen voor het herkennen van emoties (art 3 punt 39), gaat uit van een AI-systeem dat is bedoeld voor het vaststellen of afleiden van de emoties of intenties van natuurlijke personen op basis van hun biometrische gegevens. Het gaat om emoties identificeren en afleiden. Belangrijk aspect is dus dat het verbod op emotieherkenning is beperkt tot de afleiding van emoties op basis van de biometrische gegevens van een persoon. Het verbod geldt niet als emotieherkenning plaatsvindt voor medische of veiligheidsdoeleinden.

Het verbod kan overigens niet worden opgeheven door bijvoorbeeld toestemming te vragen.

Voorbeelden van toegestane AI in het onderwijs:

- Een AI-systeem dat emoties afleidt uit geschreven tekst om de stijl of toon van een bepaalde tekst te bepalen, is niet gebaseerd op biometrische gegevens en valt niet onder het verbod.
- Emotieherkenning kan uit medische overwegingen worden ingezet om leerlingen met autisme te helpen of om de toegankelijkheid voor blinden of doven te verbeteren (uitzondering van artikel 5, lid 1, punt f AI-Verordening).
- Het opnemen en met AI laten analyseren van opnames van gegeven lessen en interactie met leerlingen door LIO's/docenten voor video-interactietraining is toegestaan mits hierbij geen emotieherkenning plaatsvindt (en aan overige eisen zoals AVG is voldaan).
- Indien emotieherkenning alleen wordt gebruikt voor leerdoeleinden in het kader van een rollenspel (bijvoorbeeld om acteurs of leerkrachten op te leiden), zijn systemen voor emotieherkenning toegestaan als de resultaten niet van invloed zijn op de beoordeling of certificering van de persoon die wordt opgeleid.
- Een AI-toepassing die emotieherkenning gebruikt voor het leren van een taal *buiten de schoolomgeving* (bijvoorbeeld gekocht door ouders) is niet verboden, zolang de school het gebruik niet verplicht. Hierbij geldt dat er altijd een individuele beoordeling moet plaatsvinden: is de leerling feitelijk vrij om de AI-toepassing (niet) te gebruiken of ondervindt deze daar toch nadeel van tijdens het volgen van de les omdat de leerkracht hier toch dagelijks mee oefent.
- Het opnemen van gesprekken tussen leerlingen tijdens online lessen is niet verboden, zolang er geen emoties worden geanalyseerd (en daarnaast aan bijv. AVG-eisen is voldaan).
- Het gebruik van AI voor oogtracking of gezichtsherkenning tijdens examens (hoog-risico-AI-proctoring) is toegestaan als het systeem geen emoties identificeert of afleidt. Het volgen van het "gaze point" (waar iemand naar kijkt)

om gebruik van ongeoorloofd materiaal te detecteren is toegestaan maar het detecteren van zenuwachtigheid of stress is dus verboden.

Voorbeelden verboden AI in het onderwijs:

- Een AI-systeem dat emoties afleidt uit de toetsaanslag (manier van typen), gezichtsuitdrukkingen, lichaamshoudingen of bewegingen is gebaseerd op biometrische gegevens en valt onder het verbod.
- Een AI-systeem dat tijdens toetsing emoties meet (bijv. angst of opwinding) met behulp van biometrie, is verboden.
- Het gebruik van emotieherkenning om de aandacht of interesse van leerlingen te meten is altijd verboden in onderwijscontexten.
- Emotieherkenning tijdens toelatingstests voor nieuwe leerlingen is niet toegestaan, ongeacht het doel.
- Het inzetten van emotieherkenning voor zowel leraren (op de werkplek) als leerlingen (in het onderwijs) is verboden.

3. Is er sprake van een hoog-risico-AI-systeem?

Hoog-risico-AI-systemen zijn systemen die significante risico's opleveren voor de gezondheid, veiligheid en fundamentele rechten van personen (art. 6 lid 2 verwijst naar Bijlage 3). De AI-Verordening stelt aan deze systemen zwaardere eisen, waaronder een conformiteitsbeoordeling, een risicobeheerssysteem, transparantieplichtingen en menselijk toezicht. Het gaat altijd om de individuele afweging van de context, en de gevolgen en impact voor de leerling.

Of een AI-systeem als hoog risico wordt gekwalificeerd, hangt af van het beoogde gebruik. Voor een aantal situaties beschrijft de AI-Verordening wanneer AI-systemen een hoog risico hebben, deze situaties zijn opgenomen in Bijlage 3 van de AI-Verordening. Voor het onderwijs zijn deze toepassingsgebieden (use cases) opgenomen in onderdeel 3 (onderwijs en beroepsopleiding).

Er kunnen specifieke gevallen zijn waarin de in Bijlage III genoemde AI-systemen geen significant risico vormen voor de gezondheid, veiligheid en fundamentele rechten, omdat ze de besluitvorming niet wezenlijk beïnvloeden of die belangen niet wezenlijk schaden. Op basis hiervan zijn vier uitzonderingen genoemd waarbij er geen sprake is van een hoog-risico-AI-systeem. Deze uitzonderingen zijn opgenomen na de beschrijving van de hoog-risico-toepassingsgebieden.

Een leverancier kan niet zomaar een AI-systeem als "niet-hoog-risico classificeren" door het beoogde doel in te perken of buiten de definitie van hoog risico te brengen. De werkelijke functie, het gebruik en de impact van het AI-systeem bepalen de risicoklasse. Daarbij moet de leverancier ook rekening houden met eventueel te voorzien misbruik (bijv. discriminatie of verkeerd gebruik). Het gaat altijd om een case-by-case beoordeling.

Een belangrijk detail is dat een AI-toepassing allereerst onder de definitie "AI-systeem" (stap 1) moet vallen en geen verboden AI-praktijk is (stap 2), wil er sprake kunnen zijn van hoog risico (stap 3). Hierna wordt afgewogen of er sprake is van een uitzondering.

Voor het onderwijs zijn er vier specifieke use-cases gedefinieerd (Bijlage III, artikel 3):

- Systemen voor toelating tot of toewijzing van (beroeps)onderwijs: AI-systemen die bedoeld zijn om toegang of toelating te bepalen of om natuurlijke personen toe te wijzen aan instellingen voor onderwijs en beroepsopleiding op alle niveaus.
- Systemen voor het evalueren van leerresultaten: AI-systemen die bedoeld zijn om leerresultaten te evalueren, ook wanneer die resultaten worden gebruikt om het leerproces van natuurlijke personen in instellingen voor onderwijs en beroepsopleiding op alle niveaus te sturen.

- Systemen voor het beoordelen van het (beroeps)onderwijsniveau: AI-systemen die bedoeld zijn om te beoordelen welk onderwijsniveau een persoon zal volgen of kan volgen, in de context van of binnen instellingen voor onderwijs en beroepsopleiding op alle niveaus.
- Systemen voor het monitoren van deelnemers tijdens examens: AI-systemen die bedoeld zijn om verboden gedrag van leerlingen tijdens toetsen te monitoren en op te sporen, in de context van of binnen instellingen voor onderwijs en beroepsopleiding op alle niveaus.

In de volgende paragrafen wordt elk van deze risicocategorieën verder toegelicht.

A. Toegang tot onderwijsinstellingen

AI-systemen die bedoeld zijn om te worden gebruikt voor het bepalen van toegang of toelating tot of het toewijzen van natuurlijke personen aan instellingen voor onderwijs en beroepsonderwijs op alle niveaus; (Bijlage III, 3(a)).

Hulpvragen:

- Neemt het systeem (mede) de beslissing wie toegelaten wordt?
 - Het AI-systeem speelt een substantiële rol in de besluitvorming (maar hoeft niet de enige beslisser te zijn).
 - Er is bij geautomatiseerde besluitvorming mogelijk sprake van samenloop met artikel 22 AVG.
- Heeft het systeem directe invloed op wie wel/niet geplaatst wordt?
- Speelt het AI-systeem een substantiële rol in de besluitvorming (bijv. door een ranking, advies of voorspelling)?

Voorbeelden hoog risico:

- Geautomatiseerde selectie bij toelating tot een school.
- AI-ondersteunde lotingssystemen voor aanmelding waarbij het systeem invloed heeft op de selectiecriteria of kansverdeling.
- Automatische toewijzing aan opleidingsniveaus of tracks.

Voorbeelden géén hoog risico:

- Oriëntatietest die leerlingen helpt bij schoolkeuze zonder invloed op toelating.
- Lotingssysteem bij aanmelding en inschrijving (willekeurig in uitkomst en zonder enige vorm van redenering of patroonherkenning).

B. Evaluatie van leerresultaten

AI-systemen die bedoeld zijn om te worden gebruikt voor het evalueren van leerresultaten, ook wanneer die resultaten worden gebruikt voor het sturen van het leerproces van natuurlijke personen in instellingen voor onderwijs en beroepsonderwijs op alle niveaus (Bijlage III, 3(b)).

Deze evaluatie heeft impact op cijfers, rapporten of beslissingen en het sturen van het leerproces (niveau, tempo, content, interventies). Ook bij deze categorie is van belang per geval af te wegen of er gevolgen zijn voor de natuurlijk persoon en wat de impact daarvan is. Des te meer of groter de (mogelijke) impact, des te sneller is sprake van deze categorie hoog-risico-AI.

Hulpvragen:

- Is er sprake van een “definitieve” beslissingen die van invloed (kunnen) zijn op het onderwijs- en beroepstraject (summatieve toetsing of beoordeling)?
- Biedt het AI-systeem feedback en leerondersteuning en geeft het in dat kader sturing aan het leerproces, en is de leerling in grote mate afhankelijk van deze sturing (des te groter en meer invloed de sturing heeft, des te sneller er sprake is van hoog risico)?
- Stuurt of ondersteunt het systeem het schoolwerk en leerproces (bijv. niveau, tempo, interventies) of dient het ter ondersteuning van de leerling of docent?
- Worden de leerresultaten gebruikt om het leerproces te sturen?
- Heeft de evaluatie significante invloed op het onderwijspad van de leerling?

Voorbeelden hoog risico:

- Automatische beoordeling van toetsen met invloed op resultaten/cijfers.
- AI-systeem dat spreekvaardigheid of leesvaardigheid beoordeelt, leesniveau bepaalt en/of (eind)cijfers daarvoor geeft.
- Automated essay scoring voor eindwerkstukken.
- Summatieve toetsing indien bijdraagt aan een beoordeling, toetsing of beslissing of gevolgen voor de leerling (bijv. geautomatiseerd nakijken).

Voorbeelden géén hoog risico:

- Adaptieve leersystemen als onderdeel van het doorlopende onderwijsproces, die de moeilijkheidsgraad van cursusmateriaal aanpassen op basis van de prestaties

van een student zonder dat dit systeem van wezenlijke invloed is op de eindbeoordeling van de leerlingen. Uitzonderingen:

- een adaptief AI-systeem past zich aan de leerling aan en heeft daarmee invloed, maar het gaat om wezenlijke, significante, aanzienlijke invloed hebben op het onderwijstraject van een leerling;
- tenzij hier wel direct definitieve beslissingen over de leerling uit volgen;
- de moeilijkheidsgraad (niveau, tempo, content, interventies) niet tussentijds wordt bijgesteld of gereset. Als een adaptief AI-systeem iedere week reset, en iedereen laat starten op hetzelfde niveau, dan is de impact beperkter dan wanneer het AI-systeem pas na een jaar het niveau terugzet;
- dit adaptieve leersysteem de enige bron is die de leerling gebruikt bij een individueel vak, zodat de invloed van het AI-systeem op het leren door leerling groot zal zijn.
- AI-chatbot die uitleg geeft over lessen, huiswerk of opdrachten maar geen prestaties beoordeelt.
- Dashboard dat leerdata toont zonder evaluatieve conclusies.
- AI-systemen die door leerlingen/ouders thuis worden ingezet, behoren niet tot formeel onderwijs en/of worden niet ingezet door onderwijsinstellingen zelf. Het gaat om ondersteuning van niet-formeel onderwijs en AI-systemen die niet leiden tot het behalen van een diploma of certificaat en/of geen aanzienlijke invloed kunnen hebben op het onderwijs- en beroepstraject van leerlingen (denk aan huiswerkbegeleidingssoftware, software die door ouders wordt gebruikt voor hun kinderen).
- AI-systemen met formatieve toetsing, beoordeling of feedback zonder consequenties (bijv. oefentoetsen met feedback, maar geen cijfers) tenzij deze systemen wezenlijke invloed hebben op het onderwijs- en beroepstraject of die leiden tot eindbeoordelingen.

Kanttekening bij voorbeelden: des te groter de mogelijke gevolgen en impact kunnen zijn voor de leerling, des te sneller er toch sprake kan zijn van hoog-risico-AI. Centraal staat de significante impact op leerprestaties of leerproces. Als een leerling gedurende zijn 6-jarige vo-schoolloopbaan voor één vak gebruik maakt van één AI-adaptief leersysteem (met enkel formatieve toetsing), dan zal dat zeker van invloed zijn op het leerproces, geleerde ervaring en expertise. En daarmee op de te behalen resultaten.

Let op: een AI-systeem dat door de leverancier is ontwikkeld met beoogd doel huiswerkondersteuning (geen hoog risico), en dat door een docent wordt ingezet voor summatieve toetsing en enige bron van kennis en huiswerk voor de leerling, wordt door dit (niet door de leverancier bedoelde) gebruik wél aangemerkt als hoog risico.

Formatieve evaluatie zoals oefentoetsen met feedback is doorgaans géén hoog risico, tenzij het structurele gevolgen heeft of kan hebben voor het leerproces (bijv. aanpassing van het leerniveau). Het is een continu proces waarbij het leren en de voortgang van studenten wordt beoordeeld, met als doel de prestaties te verbeteren.

Summative evaluatie verwijst naar de beoordeling van het leren van leerlingen aan het 'einde' van een onderwijstraject (cursus, proefwerk, programma of opleidingstraject), wat leidt tot een eindcijfer, diploma, evaluatie of kwalificatie, en dus een aanzienlijke invloed kan hebben op het onderwijstraject van de leerling.

Let op: het is dus niet zo dat per definitie summatief toetsen 'altijd' hoog-risico-AI is, en formatief toetsen niet; het gaat altijd om de individuele afweging van de context, en de gevolgen en impact voor de leerling.

De AI Office van de Europese Commissie overwoog in 2025⁸: *“Formatieve evaluatie is een essentieel onderdeel van het onderwijsproces dat met behulp van digitale technologieën, waaronder AI, een grotere personalisering van het proces mogelijk maakt. Volgens de diensten van de Commissie wordt formatieve evaluatie opgevat als een doorlopend beoordelingsproces dat het onderwijstraject begeleidt, in plaats van te leiden tot een eindcijfer, evaluatie of kwalificatie. Het primaire doel ervan is het leren te ondersteunen door middel van activiteiten zoals oefeningen, feedback en begeleiding om studenten te helpen vooruitgang te boeken. Gezien de impact op het onderwijstraject van de leerling, is het voorlopige standpunt van de diensten van de Commissie dat [evaluatie van leerresultaten] voornamelijk betrekking heeft op AI-systemen die een evaluatie maken met het oog op “definitieve” beslissingen die van invloed zijn op het onderwijs- en beroepstraject, op basis van summatieve beoordeling, en niet op AI-systemen die feedback en leerondersteuning bieden als onderdeel van het lopende onderwijsproces. Dit betekent dat AI-systemen die een aanzienlijke invloed hebben op het onderwijstraject van een student, zoals AI-gestuurde beoordelingssystemen die het eindcijfer of de academische status van een student bepalen, als risicovol zouden worden beschouwd.”*

C. Bepalen van onderwijsniveau

AI-systemen die bedoeld zijn om te worden gebruikt voor het beoordelen van het passende onderwijsniveau dat een persoon zal ontvangen of waartoe hij toegang zal hebben, in het kader van of binnen instellingen voor onderwijs en beroepsonderwijs op alle niveaus (Bijlage III, 3(c)).

⁸ Vertaald naar het Nederlands uit: *The AI Board Workshop on High-Risk Classification – Education and vocational training – Annex III Point 3*

Het AI-systeem beoordeelt op welk onderwijsniveau een leerling past. Deze beoordeling heeft invloed op plaatsing in een niveau, track of leerweg en heeft impact op de onderwijsloopbaan.

Hulpvragen:

- Beoordeelt het systeem op welk onderwijsniveau een leerling past?
- Wordt het AI-systeem gebruikt om te bepalen of een student in aanmerking komt voor een bepaald onderwijsprogramma, bepaalde opleiding of een bepaald studieniveau, of om een specifiek leertraject of onderwijstraject aan te bevelen?
- Heeft deze beoordeling invloed op plaatsing in niveau/track/leerweg?
- Is dit een structurele niveaubepaling (niet tijdelijk)?
- Heeft dit (op langere termijn) impact op de onderwijsloopbaan?

Voorbeelden hoog risico:

- Systemen die schooladvies genereren of beïnvloeden.
- Systemen die adviseren over ondersteuning of advisering passend onderwijs.
- AI die leerprestaties beoordeelt voor bevorderingsbeslissingen.
- AI-aangestuurde adaptieve plaatsingstool om het optimale cursusniveau voor nieuwe leerlingen te bepalen.
- Over het algemeen geldt dat des meer belang of gewicht wordt toegekend aan de AI-systeem toetsing, test of assessment, des te sneller er sprake kan zijn van hoog-risico-AI. Voorbeeld is de doorstroomtoets waar wettelijke gevolgen aan zijn verbonden.

Voorbeelden géén hoog risico:

- Kortstondige niveau-aanpassingen in oefensoftware.
- Keuze zonder niveaucomponent (bijv. keuzevakken).
- Digitaal leermateriaal dat een signaal geeft aan de docent als de leerling laag constant slecht presteert.
- Tijdelijke groepsindeling binnen een klas.

De AI Office van de Europese Commissie overwoog in 2025⁹: “[bepalen van onderwijsniveau] heeft betrekking op AI-systemen die bedoeld zijn om te beoordelen welk

⁹ Vertaald naar het Nederlands uit: *The AI Board Workshop on High-Risk Classification – Education and vocational training – Annex III Point 3*

onderwijsniveau een persoon zal volgen of kan volgen, in de context van of binnen instellingen voor onderwijs en beroepsopleiding op alle niveaus, kunnen de volgende stappen in het leertraject van een student bepalen. Deze systemen kunnen bijvoorbeeld uitgaan van het huidige kennis- en vaardigheidsniveau van de leerling op een bepaald vakgebied, zoals wiskunde of talen, en op die manier het meest geschikte onderwijsniveau bepalen dat hij of zij vervolgens kan volgen. Dit kan onder meer inhouden dat wordt vastgesteld of de leerling kan doorgaan naar het volgende niveau van zijn of haar onderwijsproces. AI-tools om specifieke cursussen, onderwerpen of leermaterialen aan te bevelen waarmee de student zich zou moeten bezighouden, op basis van zijn individuele behoeften en vaardigheden. Deze systemen kunnen ook rekening houden met eventuele speciale behoeften of talenten van een student, zoals dyslexie of een buitengewoon muzikaal talent, en gepersonaliseerde aanbevelingen doen voor zijn opleiding. Bijvoorbeeld, een leerling met een bijzonder talent voor wiskunde kan worden aanbevolen om gevorderde cursussen in calculus of getaltheorie te volgen, terwijl een leerling met dyslexie kan worden aanbevolen om extra ondersteuning te krijgen bij lezen en schrijven.”.

D. Monitoring en detectie van ongeoorloofd gedrag tijdens toetsen

AI-systemen die bedoeld zijn om te worden gebruikt voor het monitoren en detecteren van ongeoorloofd gedrag van studenten tijdens toetsen in de context van of binnen instellingen voor onderwijs en beroepsonderwijs op alle niveaus (Bijlage III, 3(d)).

Bij deze categorie monitort het AI-systeem het gedrag van leerlingen tijdens toetsen/tentamens/examens. Het systeem gebruikt AI om verboden gedrag (fraude, spieken) te detecteren. Overigens is emotieherkenning met biometrische gegevens bij dit hoog risico ook verboden (verboden AI-praktijk).

Hulpvragen:

- Monitort het systeem gedrag van leerlingen/studenten?
- Gebeurt dit tijdens een toets/tentamen/examen?
- Gebruikt het systeem AI om verboden gedrag (fraude, spieken) te detecteren?
- Kunnen (maar niet verplicht) bevindingen leiden tot sancties?

Voorbeelden hoog risico:

- AI-systeem dat via webcam oogbewegingen analyseert tijdens een online tentamen.
- Anti-spieksoftware met gezichtsherkenning of gedragsanalyse.
- Systemen die "verdacht gedrag" detecteren via webcam/microfoon.

Voorbeelden géén hoog risico:

- Plagiaatdetectie van ingeleverde werkstukken na afronding, omdat dit niet valt onder de definitie "tijdens toetsen", zo legt de Europese Commissie¹⁰ uit. Mogelijk heeft dit AI-systeem wél significante gevolgen voor leerlingen zodat dit valt onder het hoog risico b) evaluatie leerresultaten.
- Camera-opname of cameratoezicht zonder AI-analyse (alleen menselijke beoordeling).
- Administratief systeem dat registreert wie tentamen heeft afgelegd (geen gedragsanalyse).

Toelichting:

- Het systeem moet realtime of direct na afloop gedrag monitoren, dus tijdens de toetsing.
- Let op verbod op emotieherkenning: een camera kan wel monitoren waar iemand bijvoorbeeld naar toe kijkt zonder emoties te herkennen. Maar zenuwachtig gedrag herkennen (verdenking spieken) valt mogelijk onder het verbod op emotieherkenning.

Overige voor onderwijs relevante hoog-risico-AI-systemen

Bij deze overige hoog-risicotoepassingen betreft het AI-systemen die door de school als gebruiksverantwoordelijke worden ingezet. Dus niet systemen die andere gebruikersverantwoordelijken (buiten de school-context) gebruiken voor leerlingen. Denk bijvoorbeeld aan gemeente (jeugdgezondheidszorg) of DUO/studiefinanciering.

¹⁰ Vertaald naar het Nederland uit: *The AI Board Workshop on High-Risk Classification – Education and vocational training – Annex III Point 3*

Biometrie en Biometrische Identificatie

AI-systemen die biometrische gegevens verwerken voor identificatie of categorisering vallen onder hoog risico als ze significante risico's opleveren. (Bijlage III, punt 1, categorie 3)

Voorbeelden:

- Biometrische identificatie op afstand (niet realtime in openbare ruimte):
- Achteraf gezichtsherkenning in camerabeelden (bijv. voor veiligheidsdoeleinden).
- Biometrische categoriseringssystemen (niet onder verbod Artikel 5):
- Systemen die leeftijd of geslacht inschatten (niet voor afleiden ras/religie).
- Emotieherkenning in context anders dan onderwijs/werkplek, of voor medische/veiligheidsdoeleinden op school.

Toegang tot Essentiële Diensten

AI-systemen die worden gebruikt voor het beoordelen van toegang tot essentiële diensten vallen onder hoog risico. Dit kan relevant zijn voor scholen als ze AI-systemen gebruiken om leerlingen te beoordelen voor toegang tot sociale diensten. (Bijlage III, punt 5, categorie 7)

Voorbeelden:

- AI-systemen bedoeld om door overheidsinstanties of namens hen te worden gebruikt voor het beoordelen van geschiktheid van natuurlijke personen voor overheidsdiensten (bijv. sociale bijstand, kinderopvang, gezinsondersteuning).
- AI-systemen voor kredietwaardigheid of credit scoring (behalve voor fraudedetectie).
- AI-systemen voor risicobeoordeling en prijsberekening in levens- en ziektekostenverzekeringen.
- AI-systemen voor dispatching of vaststelling van prioriteit bij hulpdiensten (112, ambulance, brandweer, politie).

Werkgelegenheid en Werknemersbeheer

AI-systemen die worden gebruikt voor werving, selectie, beoordeling en monitoring van werknemers vallen onder hoog risico. Dit is relevant voor scholen die AI-systemen gebruiken voor HR-doeleinden. (Bijlage III, punt 5, categorie 6)

Voorbeelden:

- AI-systemen bedoeld voor werving en selectie van personeel.
- AI-systemen voor het nemen van beslissingen die van invloed zijn op arbeidsvoorwaarden, promotie en beëindiging van contractuele relaties.
- AI-systemen voor het toewijzen van taken op basis van individueel gedrag, persoonlijke kenmerken of eigenschappen.
- AI-systemen voor het monitoren en beoordelen van prestaties en gedrag van personeel.

Rechtshandhaving en Individuele Risicobeoordeling

De categorie rechtshandhaving zal zelden relevant voor onderwijs zijn tenzij het gaat om onderwijs in justitiële setting voor veiligheidsdoeleinden of veiligheidsprogramma's. Het gaat om AI-systemen die worden gebruikt voor individuele risicobeoordeling om het risico op criminele activiteit te beoordelen. (Bijlage III, punt 6, categorie 8(i))

Voorbeelden:

- AI-systemen bedoeld voor individuele risicobeoordeling voor natuurlijke personen om risico op criminele activiteit te beoordelen.
- AI-systemen bedoeld als polygraaf of soortgelijke tool voor detectie van emotionele staat.
- AI-systemen bedoeld voor detectie van deep fakes.
- AI-systemen bedoeld voor beoordeling van betrouwbaarheid van bewijs in strafonderzoek.
- AI-systemen bedoeld voor beoordeling van risico op recidive of re-integratie.
- AI-systemen bedoeld voor profilering in opsporingsonderzoek.
- AI-systemen bedoeld voor forensische analyse (spraak, handschrift, etc.).

Uitzonderingen hoog-risico-AI

Een AI-systeem kwalificeert niet als AI-systeem met een hoog risico als het geen significant risico op schade voor de gezondheid, veiligheid of de grondrechten van natuurlijke personen inhoudt, onder meer doordat het de uitkomst van de besluitvorming niet wezenlijk beïnvloedt. Een AI-systeem dat het resultaat van de besluitvorming niet wezenlijk beïnvloedt, wordt beschouwd als een AI-systeem dat geen invloed heeft op de inhoud, en daarmee op het resultaat, van de menselijke of geautomatiseerde besluitvorming.

Bij de vaststelling dat er mogelijk sprake is van een AI-systeem met hoog risico (genoemd in Bijlage 3), kan er sprake zijn van een van de vier uitzonderingen op de hoog-risico-classificatie. Deze uitzonderingen gelden niet als er sprake is van profilering (art 6 lid 3 laatste volzin AI-Verordening).

Er is geen sprake van een AI-systeem met hoog risico als het AI-systeem:

- a) een beperkte procedurele taak uitvoert (bijv. administratieve taken zonder invloed op besluitvorming zoals een AI-systeem dat ongestructureerde gegevens omzet in gestructureerde gegevens, een AI-systeem dat binnenkomende documenten in categorieën indeelt of een AI-systeem dat wordt gebruikt om duplicaten te detecteren in een groot aantal aanvragen: deze taken zijn zo beperkt (in aard en omvang) dat ze slechts beperkte risico's met zich meebrengen).
- b) een eerder voltooide menselijke activiteit verbetert (bijv. een AI die een door een docent gemaakte beoordeling verfijnt, zonder de uitkomst te veranderen zoals AI-systemen om de taal in eerder opgestelde documenten te verbeteren met betrekking tot de professionele toon, de academische stijl of door de tekst af te stemmen op een bepaalde doelgroep).
- c) is bedoeld om besluitvormingspatronen of afwijkingen van eerdere besluitvormingspatronen op te sporen en niet is bedoeld om de eerder voltooide menselijke beoordeling zonder behoorlijke menselijke toetsing te vervangen of te beïnvloeden. Het gaat bijvoorbeeld om systemen die - gegeven een bepaald beoordelingspatroon van een leraar - kunnen worden gebruikt om achteraf te controleren of de leraar mogelijk is afgeweken van het eerdere beoordelingspatroon met als doel mogelijke inconsistenties of anomalieën te signaleren. De docent blijft wel verantwoordelijk voor de uiteindelijke beoordeling.
- d) is bedoeld om een voorbereidende taak uit te voeren voor een beoordeling die relevant is voor de in bijlage III (4 hoge risico-categorieën) vermelde gebruiksgevallen. Bijvoorbeeld een AI-systeem dat gegevens voorbereidt voor een menselijke beoordeling, zonder zelf beslissingen te nemen zoals slimme oplossingen voor bestandsverwerking waaronder indexeren, zoeken, tekst- en spraakverwerking of het koppelen van gegevens aan andere gegevensbronnen, of AI-systemen die worden gebruikt voor de vertaling van oorspronkelijke documenten).

Bij deze uitzonderingen op de hoog-risicocategorie is menselijke tussenkomst alleen relevant als die tussenkomst daadwerkelijk betekenisvol is. Van betekenisvolle menselijke beoordeling is geen sprake wanneer een medewerker de uitkomst van het AI-systeem in de praktijk blind of routinematig overneemt (over kan nemen) zonder

eigen inhoudelijke toetsing. Deze zogenoemde *rubber stamping* moet worden voorkomen. Als het AI-systeem feitelijk wél de menselijke beoordeling vervangt, stuurt of beïnvloedt, dan kan er geen beroep worden gedaan op de uitzonderingen.

Profilering

AI-systemen uit Bijlage III van de AI-Verordening die gebruikmaken van profilering (het analyseren van persoonlijke gegevens om voorspellingen te doen over gedrag, voorkeuren of capaciteiten) zijn altijd hoog risico. De uitzonderingen zijn dan niet van toepassing. Bij een dashboard dat (statisch) cijfers registreert en leerresultaten evalueert en weergeeft, is geen sprake van profilering terwijl bij een dashboard dat analyses en voorspellingen uitvoert dat wel kan zijn. Bij een AI-systeem zoals een dashboard, leerlingadministratiesysteem (LAS) of leerlingvolgsysteem (LVS) dat persoonlijke gegevens analyseert en voorspellingen doet over (individueel) gedrag, voorkeuren of capaciteiten, is sprake van profilering (22 AVG) zodat een beroep op de uitzondering niet mogelijk is en er daarmee sprake is van een AI-systeem met hoog risico.

Als er bij een AI-systeem profilering wordt toegepast, betekent dat overigens niet dat dit AI-systeem hoog risico is. Profilering is de uitzondering op de uitzondering voor hoog risico.

Verschuiving van aanbiedersverantwoordelijkheid

Het is belangrijk dat scholen AI-systemen gebruiken volgens de instructies van de leveranciers. Leveranciers hebben hun ict-systemen zo ingericht dat bij gebruik volgens de instructie de risico's beperkt worden (blijven). Als scholen iets anders doen met AI dan waarvoor het ontworpen is, dan kan dat grote gevolgen hebben op de levens van leerlingen en medewerkers als zij bijvoorbeeld gediscrimineerd of uitgesloten worden.

Het volgen van de instructies van de leveranciers, is helemaal belangrijk bij AI-systemen met hoog risico. Van belang is dat scholen AI-systemen niet inzetten op een manier waarvoor het systeem niet bedoeld is. Algemene (generatieve) AI-modellen (ChatGPT, Gemini, Copilot, Claude, Mistral, etc) zijn niet specifiek ontworpen voor onderwijs en dat kan risico's met zich meebrengen. Denk bijvoorbeeld aan gebrek aan pedagogische afstemming, privacyproblemen of onvoldoende waarborgen voor veilig en rechtvaardig gebruik. Ook zijn deze AI-systemen niet ingericht op het verwerken van persoonsgegevens van minderjarige leerlingen.

Educatieve AI (bijv. adaptieve leerplatforms), waar AI (al dan niet met behulp van generieke AI-modellen) in educatieve applicaties is ingebouwd, is veiliger en doelmatiger omdat deze systemen wel zijn afgestemd op onderwijscontexten en vaak betere bescherming bieden.

Het is belangrijk om bij de inzet van AI op school, dat docenten leerlingen een toereikend niveau van AI-geletterdheid bij hebben, en op de hoogte zijn van mogelijke nadelige gevolgen en risico's. Als docenten op scholen AI-modellen, AI-systemen of generatieve AI (zoals Copilot, Claude of ChatGPT) aanpassen, bijvoorbeeld via een systeem-prompt, fine-tuning of geavanceerde inrichting of integratie inricht voor een specifieke toepassing, kan die toepassing van AI leiden tot een hoog risico. Datzelfde geldt als docenten vibecoden op een platform dat niet is ingericht voor bijvoorbeeld minderjarigen of onderwijs. Dat kan leiden tot risico op schade voor de gezondheid, veiligheid of de grondrechten van de leerlingen of medewerkers. Voorbeeld: een docent die ChatGPT ombouwt tot een nakijktol, en de tool beoordeelt of de antwoorden goed of fout zijn en een cijfer geeft, dan is er sprake van 'beoordelen leerresultaten': hoog-risicocategorie uit Bijlage III.

Het gebruiken van AI-systemen voor een ander doel dan het beoogde doel kan ertoe leiden dat volgens artikel 25 de rol van aanbieder verschuift van leverancier naar school. Een school is in beginsel gebruiksverantwoordelijke en de leverancier aanbieder. De school kan echter juridisch als aanbieder worden aangemerkt indien:

- de school een reeds bestaand hoog-risico-AI-systeem onder eigen naam of merk op de markt brengt of
- in gebruik stelt, of
- daarop een substantiële wijziging aanbrengt, of
- het beoogde doel van een bestaand AI-systeem — waaronder ook een general-purpose AI-systeem — zodanig wijzigt dat het daardoor een hoog-risico-AI-systeem wordt.

Met name dit laatste is het geval als docenten generatieve AI inzetten voor hoog-risico-toepassingen. Het AI-systeem is daarvoor niet gebouwd en bedoeld. Omdat de school nu gezien wordt als aanbieder, moet de school aan de volledige set aanbiedersverplichtingen gaan voldoen. Dat zal praktisch onmogelijk zijn omdat de school bijvoorbeeld geen toegang heeft tot de (bron)code.

Het kantelpunt ligt daar waar de school het AI-systeem wezenlijk wijzigt of functioneel aanpast. Bijvoorbeeld wanneer een docent een generiek AI-systeem zoals ChatGPT of Copilot niet gebruikt als ondersteunend hulpmiddel, maar buiten het oorspronkelijke doel doorontwikkelt, inzet of feitelijk gebruikt. Denk bijvoorbeeld aan het beoordelen van leerresultaten, huiswerk of toetsen laten beoordelen, het leerproces sturen, het

onderwijsniveau van leerlingen bepalen of verboden gedrag tijdens toetsen monitoren. Er kan dan sprake zijn van een hoog-risicotoepassing waarvoor de leverancier aanvullende maatregelen moet nemen. En omdat de school nu aanbieder is, moet de school die maatregelen nemen. Volgens artikel 25 geldt de oorspronkelijke aanbieder van dat specifieke AI-systeem dan niet langer als aanbieder. Het gaat dus niet om regulier gebruik van een bestaand AI-systeem overeenkomstig het door de leverancier omschreven beoogde doel. Artikel 25 is alleen van toepassing als de gebruiksverantwoordelijke (of importeur of distributeur) een AI-systeem inzet voor hoog-risicotoepassing.

Voorbeeld: als een docent van school zelf ChatGPT of Copilot ombouwt tot een volledig nakijk-, correctie- en beoordelingssysteem, dan wordt deze GPAI niet gebruikt waarvoor het bedoeld is. De school wordt dan gezien als leverancier (aanbieder) van dat systeem en moet zelf aan de leveranciersverplichtingen voldoen zoals een conformiteitsbeoordeling bij hoog-risico-AI.

Gezien de gevolgen van artikel 25 voor scholen, is het belangrijk om op school de instructie te geven AI-systemen alleen in te zetten volgens de instructies van de leverancier, en om afwijking daarvan niet toe te staan (om te voorkomen dat AI ingezet wordt voor hoog-risicotoepassing zonder dat de school daar zich op heeft). Daarnaast is het - gezien de vele maatregelen die een school al als gebruiksverantwoordelijke moet nemen - het advies om niet door de school (docenten) zelf-ontwikkelde AI-systemen met hoog risico in te zetten.

4. Maatregelen voor hoog-risico-AI-systemen

Als er sprake is van een hoog-risico-AI-systeem, dan moeten de leverancier en de school een aantal maatregelen nemen. Deze maatregelen vallen uiteen in drie categorieën;

1. eisen aan het systeem,
2. verplichtingen voor de leverancier,
3. verplichtingen voor de school.

Er kan hierbij sprake zijn van een wisselwerking. De eisen waaraan het systeem moet voldoen en de verplichtingen die de gebruiksverantwoordelijke heeft hangen namelijk samen met de maatregelen die de leverancier moet nemen.

De AI-Verordening treedt gefaseerd in werking. Voor hoog-risico-AI-systemen als bedoeld in Bijlage III gelden de verplichtingen vanaf 2 augustus 2026. Hoog-risico-AI-systemen die vóór die datum op de markt zijn gebracht of in gebruik zijn genomen, vallen alleen onder de verordening als zij daarna ingrijpend worden gewijzigd. Voor zulke systemen die bedoeld zijn voor gebruik door overheidsinstanties geldt een overgangstermijn tot uiterlijk 31 december 2030.

Eisen voor hoog-risico-AI-systemen

Hoog-risico-AI-systemen moeten aan de volgende eisen voldoen:

- Algemene conformiteitseis (art. 8): de hoofdregel voor hoog-risico-AI-systemen is dat het systeem, gelet op het beoogde gebruik en de stand van de techniek, moet voldoen aan alle hieronder opgenomen eisen (art. 9-15). Deze naleving moet vóór het op de markt brengen of in gebruik nemen aantoonbaar zijn gemaakt via een toepasselijke conformiteitsbeoordeling (art. 43) die de leverancier uitvoert. Daarop volgend stelt de leverancier een EU-conformiteitsverklaring (art. 47) op, en brengt een CE-markering (art 48) aan.
- Risicobeheersysteem (art. 9): er moet een risicobeheersysteem zijn in de vorm van een continu proces dat de gehele levenscyclus van het AI-systeem bestrijkt, én dat regelmatig wordt geëvalueerd en bijgewerkt. De risicobeheersmaatregelen zijn erop gericht risico's weg te nemen of te beperken. Er moeten tests worden uitgevoerd om naleving en prestaties te waarborgen, waarbij rekening moet worden gehouden met personen jonger dan 18 jaar en kwetsbare groepen. Aanbieders kunnen deze vereisten integreren in bestaande risicobeheerprocessen die voortvloeien uit andere EU-wetgeving.

- Eisen aan data en data governance (art. 10): de gebruikte datasets moeten onder passende data governance en databeheer vallen. Dit betekent onder meer beoordeling van de herkomst van de data, voorbereiding, beschikbaarheid en geschiktheid van de data, onderzoek naar mogelijke bias, maatregelen om bias te voorkomen of te beperken, en herstel van datagaten of andere tekortkomingen. De datasets moeten, gelet op het beoogde gebruik, relevant, voldoende representatief en zoveel mogelijk foutvrij en volledig zijn, rekening houdend met de context van inzet. Bijzondere persoonsgegevens mogen alleen bij uitzondering en onder strikte voorwaarden worden verwerkt voor biasdetectie en biascorrectie.
- Technische documentatie (art. 11): technische documentatie moet worden opgesteld vóór het op de markt brengen en actueel worden gehouden. Bevat onder meer: beschrijving van het systeem, ontwerp- en ontwikkelingsproces, risicobeheersysteem, monitoring, validatieresultaten en conformiteitsverklaring. (bijlage IV bij de AI-Verordening bevat de minimumvereisten).
- Registratie en logging (art. 12): het hoog-risico-AI-systeem moet automatisch relevante gebeurtenissen kunnen loggen gedurende de levenscyclus. Die logging moet voldoende traceerbaarheid bieden om risico's te herkennen, post-market monitoring te ondersteunen en de werking van het systeem te controleren. Voor bepaalde biometrische systemen gelden aanvullende minimumeisen voor de inhoud van die logs.
- Transparantie en informatieverstrekking aan gebruiksverantwoordelijken (art. 13): hoog-risicosystemen moeten transparant zijn om gebruiksverantwoordelijken in staat te stellen de output van een systeem te interpreteren en op passende wijze te gebruiken en vergezeld gaan van gebruiksinstructies. De instructies moeten bevatten: identiteit van de aanbieder, kenmerken en beperkingen van het systeem, prestatieniveaus, bekende risico's, vereiste menselijk toezicht, verwachte levensduur en onderhoud. Deze transparantie-eis is daarmee inhoudelijk anders dan die van artikel 50.
- Menselijk toezicht (art. 14): hoog-risico-AI-systemen moeten effectief menselijk toezicht mogelijk maken. De gebruiksverantwoordelijke moet het systeem en de beperkingen ervan kunnen begrijpen, de werking kunnen monitoren, afwijkingen kunnen signaleren, output kunnen beoordelen en zo nodig kunnen negeren, overschrijven of het systeem veilig kunnen stoppen. Daarbij moet worden voorkomen dat blind op AI-uitkomsten wordt vertrouwd.
- Nauwkeurigheid, robuustheid en cyberbeveiliging (art 15): het systeem moet een passend niveau van nauwkeurigheid, robuustheid en cyberbeveiliging hebben gedurende de levenscyclus. Vereist: consistente prestaties, weerstand tegen

fouten/storingen/manipulatie, bescherming tegen vijandige aanvallen (adversarial attacks), en technische redundantie waar nodig.

Verplichtingen voor de aanbieder/leverancier

Aanbieders / leveranciers van AI-systemen met een hoog-risicoclassificatie moeten de volgende maatregelen nemen:

- AI Geletterdheid (art. 4): aanbieders van AI-systemen (overigens ook bij niet-hoogrisico) moeten maatregelen nemen om ervoor te zorgen dat hun personeel en andere personen die namens hen betrokken zijn bij de exploitatie en het gebruik van AI-systemen, over voldoende kennis van AI beschikken. Daarbij moeten zij rekening houden met hun technische kennis, ervaring, opleiding en training, alsook met de context waarin de AI-systemen zullen worden gebruikt en met de personen op wie de AI-systemen zullen worden toegepast.
- De aanbieder zorgt ervoor dat het hoog-risico-AI-systeem voldoet aan de vereisten van artikelen 8–15. Vóór het op de markt brengen of in gebruik nemen moet het systeem de toepasselijke conformiteitsbeoordeling hebben doorlopen.
- Algemene verplichtingen van aanbieders (art. 16: hoofdartikel): de aanbieder is eindverantwoordelijk voor conformiteit van het hoog-risico-AI-systeem. Alle verplichtingen uit artikelen 8–15 rusten primair op de aanbieder (zie opsomming hiervoor). Aanbieders vermelden op het AI-systeem met een hoog risico of, wanneer dit niet mogelijk is, op de verpakking of in de bij het product gevoegde documentatie, naargelang het geval, hun naam, geregistreerde handelsnaam of geregistreerd merk en hun contactadres.
- Kwaliteitsbeheersysteem (art. 17): om naleving van de regelgeving te waarborgen moet de aanbieder een kwaliteitsbeheersysteem opzetten en documenteren. Dit systeem heeft betrekking heeft op het ontwerp, ontwikkeling, testen, risicobeheer en gegevensverwerking. Dit beheersysteem moet procedures omvatten voor monitoring na het in de handel brengen, incidentmelding en communicatie met de autoriteiten. De invoering moet in verhouding staan tot de omvang van de aanbieder en kan worden geïntegreerd in bestaande kwaliteitsmanagementsystemen.
- Bewaring van technische documentatie (art. 18): de aanbieder bewaart de technische documentatie en de EU-conformiteitsverklaring minimaal 10 jaar na het op de markt brengen van het systeem.

- Automatisch gegenereerde logs (art. 19): aanbieders moeten automatisch gegenereerde logbestanden ten minste zes maanden bewaren (of zo lang als de toepasselijke wetgeving voorschrijft) om naleving en traceerbaarheid te waarborgen.
- Corrigerende maatregelen en informatieplicht (art. 20): aanbieders moeten onmiddellijk corrigerende maatregelen nemen als zij vermoeden/vaststellen dat niet voldaan wordt aan de verplichtingen (waaronder het uit de handel nemen, uitschakelen of terugroepen van het systeem) inclusief informeren van distributeurs, gebruikers en autoriteiten over het probleem en de genomen maatregelen. Aanbieders onderzoeken de oorzaken en werken indien nodig samen met de gebruikers en toezichthouders.
- Samenwerking bevoegde autoriteiten (art. 21): aanbieders moeten op verzoek van een toezichthouder alle benodigde informatie en documentatie verstrekken om aan te tonen dat aan de vereisten van de AI-wet wordt voldaan, en toegang verlenen tot de automatisch gegenereerde logbestanden van het systeem indien daarom wordt verzocht. Verkregen informatie wordt vertrouwelijk behandeld.
- Conformiteitsbeoordeling (art. 43): aanbieders moeten - voor het in de handel brengen of in gebruik nemen - zelf een (interne) conformiteitsbeoordeling uitvoeren (zonder een aangemelde autoriteit in te schakelen). Bij wezenlijke wijzigingen aan het AI-systeem moet een nieuwe conformiteitsbeoordelingen worden uitgevoerd (tenzij de wijzigingen vooraf zijn vastgesteld en gedocumenteerd).
- EU-conformiteitsverklaring (art. 47): De aanbieder moet voor elk hoog-risico-AI-systeem een EU-conformiteitsverklaring opstellen, bewaren gedurende tien jaar en op verzoek beschikbaar houden voor de bevoegde autoriteiten. In die verklaring moet het betreffende AI-systeem duidelijk zijn geïdentificeerd.
- CE-markering (art. 48): de aanbieder moet het hoog-risico-AI-systeem van een juiste en toegankelijke CE-markering voorzien. Die markering moet schriftelijk machineleesbaar, fysieke of elektronisch ondertekend zichtbaar en duurzaam zijn, zo nodig op verpakking of documentatie, en moet waar relevant ook het nummer van de aangemelde instantie en naleving van andere toepasselijke EU-regels vermelden.
- Registratie (art. 49): voordat AI-systemen met een hoog risico in de handel worden gebracht of in gebruik worden genomen, moeten aanbieders zichzelf en hun systemen registreren in de EU-databank voor AI-systemen met hoog risico (art 71).

- Post-market monitoring (art 72): de aanbieder stelt een post-market monitoringssysteem in dat actief gegevens verzamelt over de werking van het systeem na ingebruikname. Dit systeem maakt vroegtijdig ingrijpen mogelijk bij onverwachte risico's.
- Meldplicht ernstige incidenten (art. 73): aanbieders met een hoog risico melden ernstige incidenten bij de markttoezichtautoriteiten.

De onderstaande artikelen bevatten ketenverplichtingen: verplichtingen voor andere partijen in de AI-waardeketen die de naleving van de aanbiedersverplichtingen moeten ondersteunen, controleren en afdwingbaar maken. Deze verplichtingen zijn van toepassing op andere partijen dan de aanbieders, maar deze regels gaan over de nakoming van de verplichtingen voor de aanbieder:

- Gemachtigde vertegenwoordiger aanbieder (art. 22): aanbieders van buiten de EU moeten een gemachtigde vertegenwoordiger binnen de EU aanstellen voordat zij hun systemen op de markt brengen. De vertegenwoordiger controleert de nalevingsdocumentatie, bewaart de gegevens gedurende tien jaar en werkt samen met de autoriteiten bij verzoeken en risicobeperking.
- Importeurs (art 23): importeurs moeten nagaan of het systeem voldoet aan de AI-wet, met inbegrip van conformiteitsbeoordelingen, technische documentatie, CE-markering vergezeld van de EU-conformiteitsverklaring en de aanwijzing van een gemachtigde vertegenwoordiger. Zo niet, dan mag het AI-systeem niet op de markt worden gebracht.
- Distributeurs (art 24): distributeurs moeten controleren of het systeem voorzien is van de CE-markering, de EU-conformiteitsverklaring en de gebruiksaanwijzing, en of de aanbieders en importeurs aan hun verplichtingen hebben voldaan. Zo niet, dan mag het AI-systeem niet op de markt worden gebracht. Ook zorgen distributeurs voor de juiste opslag- en transportomstandigheden en nemen zij corrigerende maatregelen als zij na het op de markt brengen problemen constateren, waarbij zij de relevante partijen en autoriteiten op de hoogte stellen.

Verplichtingen voor de gebruiksverantwoordelijke/school

Schoolbesturen (gebruikersverantwoordelijke) die gebruik maken van een systeem met een hoog risico-classificatie moeten de volgende maatregelen nemen:

- Geletterdheid (art. 4): scholen die AI-systemen gebruiken (ook niet-hoog risico) moeten maatregelen nemen om ervoor te zorgen dat hun personeel en andere personen die namens hen betrokken zijn het gebruik van AI-systemen, over voldoende kennis van AI beschikken. Daarbij moeten zij rekening houden met

hun technische kennis, ervaring, opleiding en training, alsook met de context waarin de AI-systemen zullen worden gebruikt en met de personen op wie de AI-systemen zullen worden toegepast.

- Verplichting school als gebruiksverantwoordelijke (art. 26):
 - Gebruiksaanwijzing opvolgen: de school neemt passende technische en organisatorische maatregelen om het systeem te gebruiken in overeenstemming met de gebruiksaanwijzing die de aanbieder heeft verstrekt.
 - Menselijk toezicht organiseren: de school draagt het menselijk toezicht op aan natuurlijke personen die over de nodige bekwaamheid, opleiding en autoriteit beschikken (AI geletterd zijn) en de nodige ondersteuning krijgen.
 - Inputdata controleren: voor zover de school controle heeft over de inputdata, zorgt hij ervoor dat deze relevant en voldoende representatief zijn voor het beoogde doel van het systeem.
 - Monitoring en meldplicht bij risico's en incidenten: de school monitort de werking van het systeem op basis van de gebruiksaanwijzing. Bij risico's stelt hij de aanbieder of distributeur én de markttoezichtautoriteit (art. 72) in kennis en onderbreekt hij het gebruik. Bij een ernstig incident stelt hij onmiddellijk eerst de aanbieder in kennis, en vervolgens de markttoezichtautoriteit.
 - Bewaring van logs: de school bewaart de automatisch gegenereerde logs voor zover deze onder zijn controle vallen, gedurende een passende periode voor het beoogde doel, of ten minste zes maanden, tenzij Unie- of nationaal recht anders bepaalt.
 - Informeren van werknemers: voordat het systeem op de werkplek in gebruik wordt gesteld, informeert de gebruiksverantwoordelijke als werkgever de werknemersvertegenwoordigers en de betrokken werknemers dat zij worden blootgesteld aan het hoog-risico-AI-systeem.
 - Registratieplicht voor overheidsinstanties: gebruiksverantwoordelijken die overheidsinstanties zijn, leven de registratieverplichtingen van artikel 49 na. Indien het systeem niet in de EU-databank is geregistreerd, wordt het systeem niet gebruikt en wordt de aanbieder of distributeur daarvan in kennis gesteld.
 - DPIA uitvoeren: de school gebruikt de informatie van de leverancier over transparantie en informatieverstrekking (art. 13) om te controleren of een DPIA nodig is en voert deze ook uit (artikel 35 AVG).

- Informeren van betrokkenen: scholen moeten betrokken personen informeren over de hoog-risico-AI-systemen (van Bijlage III) die beslissingen nemen of helpen nemen over natuurlijke personen, alsmede dat die hoog-risico-AI-systemen op hen wordt toegepast.
- Medewerking aan toezichthouders: scholen werken samen met bevoegde autoriteiten bij alle maatregelen die zij nemen in het kader van de uitvoering van de verordening.
- Grondrechtenbeoordeling (FRIA; art 27): scholen voeren vóór ingebruikname een beoordeling uit van de gevolgen voor de grondrechten. Deze beoordeling omvat: beschrijving van de processen waarin het systeem wordt gebruikt, periode en frequentie van gebruik, betrokken categorieën personen, specifieke risico's op schade voor die personen, beschrijving van menselijk toezicht, en maatregelen bij het optreden van risico's. NB. SIVON ontwikkelt hiervoor dienstverlening en levert de vereiste modellen.
- Kennisgeving aan markttoezichtautoriteit (Artikel 27 lid 3): na uitvoering van de grondrechtenbeoordeling stelt de gebruiksverantwoordelijke de markttoezichtautoriteit in kennis van de resultaten, met gebruikmaking van het door het AI-bureau ontwikkelde sjabloon.
- Samenloop met DPIA (Artikel 27 lid 4): indien de grondrechtenbeoordeling overlapt met een reeds uitgevoerde DPIA op grond van artikel 35 AVG, vormt de FRIA een aanvulling daarop en hoeft het deel waarop de DPIA eerder is uitgevoerd, niet volledig opnieuw te worden uitgevoerd.
- Scholen moeten als zij kwalificeren als publiekrechtelijke instellingen (of, hoog-risico-AI-systemen registreren in de EU databank (art 71) als zij deze systemen gebruiken (Artikel 49 lid 3).
- Als de school een hoog-risico AI-systeem inzet bij besluitvorming, moet deze procedures, verantwoordelijkheden en documentatie zodanig inrichten dat aan betrokkenen op verzoek een duidelijke en betekenisvolle uitleg kan worden gegeven over de rol van het AI-systeem in de besluitvorming en de belangrijkste gronden van het genomen besluit (Artikel 86 lid 1).

Er is discussie of de school moet verifiëren of het AI-systeem een geldige CE-markering en EU-conformiteitsverklaring heeft (art. 26 lid 1) en of de aanbieder/AI-systeem zijn geregistreerd in EU-database (art. 49/71). Deze verplichting zou afgeleid zijn van de verplichting het AI-systeem te gebruiken conform de instructies. Toch is het raadplegen van deze informatie voor gebruik te adviseren omdat de school op basis van deze registraties (of het ontbreken ervan) de verplichtingen van de aanbieder/leverancier kan afleiden.

Fundamental Rights Impact Assessment (FRIA)

Een van de bovengenoemde maatregelen die een schoolbestuur moet nemen is het uitvoeren van een beoordeling van de gevolgen voor de grondrechten van AI-systemen met een hoog risico (art. 27), ook wel Fundamental Rights Impact Assessment (FRIA) genoemd. Voor het primair en voortgezet onderwijs voert SIVON deze namens de scholen uit.

Een FRIA moet specifiek aandacht besteden aan de volgende in de AI-Verordening opgenomen punten:

1. Beschrijving van het AI-systeem en de gebruiksaanpak: wat is het beoogde doel, hoe en waar wordt het systeem ingezet, en wie zijn de betrokkenen (leerlingen, docenten, ouders)?
2. Periode en frequentie van gebruik: hoe lang wordt het systeem ingezet en hoe vaak neemt het beslissingen of ondersteunt het die?
3. Betrokken personen en groepen: welke categorieën natuurlijke personen ondervinden naar verwachting gevolgen van het gebruik, en zijn er kwetsbare groepen zoals jonge kinderen of leerlingen met een beperking?
4. Identificatie van specifieke risico's: welke risico's op schade bestaan er voor de onder c) geïdentificeerde groepen? Denk aan bias, discriminatie, onjuiste beoordelingen of uitsluiting.
5. Menselijk toezicht: hoe is het toezicht ingericht overeenkomstig de gebruiksinstructies van de aanbieder? Wie kan het systeem monitoren, corrigeren of uitschakelen?
6. Governance en klachtenafhandeling: wat zijn de interne afspraken als een risico zich voordoet, en welke klachtenregeling is beschikbaar voor betrokkenen?

Er zijn modellen in ontwikkeling om te gebruiken. Als het uitvoeren van een FRIA verplicht is, en er is of wordt een DPIA uitgevoerd, dan mag de FRIA daarmee gecombineerd worden.

Transparantieverplichtingen

Transparantie is een voorwaarde om effectief menselijk toezicht op AI-systemen te kunnen uitoefenen. Bij gebruik van AI-systemen die onder de AI-Verordening vallen, moeten gebruikers goed worden geïnformeerd over het gebruik en de risico's van het systeem. Die informatie is relevant voor medewerkers, leerlingen en ouders. Artikel 50 legt daarvoor vier concrete verplichtingen op:

1. Informeren dat je met AI praat. Wie een AI-systeem inzet dat direct communiceert met mensen, denk aan een chatbot voor leerlingen of ouders,

moet die personen informeren dat zij met AI interageren. Dat geldt niet als het overduidelijk is, maar in een onderwijscontext is dat zelden vanzelfsprekend.

- Labelen van AI-gegenereerde content. Systemen die tekst, beeld, audio of video genereren moeten hun output machineleesbaar markeren als kunstmatig gegenereerd. Voor scholen die AI inzetten voor het produceren van lesmateriaal of communicatie is dit een technische verplichting die bij de aanbieder ligt, maar waarop de school als deployer moet toezien.
- Informeren bij emotieherkenning en biometrische categorisering. Als een school een systeem inzet dat emoties herkent of leerlingen biometrisch categoriseert, moeten de betrokkenen over de werking worden geïnformeerd. Verwerking van persoonsgegevens moet daarnaast voldoen aan de AVG, deze twee verplichtingen lopen dus samen.
- Openbaarmaking van deepfakes en AI-tekst voor publieke informatie. Wie AI-gegenereerde of bewerkte content publiceert om het publiek te informeren, bijvoorbeeld via de schoolwebsite of nieuwsbrief, moet duidelijk maken dat die content kunstmatig is gegenereerd. Een uitzondering geldt als de tekst redactionele controle heeft ondergaan en een natuurlijke of rechtspersoon daarvoor verantwoordelijkheid draagt.

De informatieplicht geldt steeds uiterlijk bij het eerste contact of de eerste blootstelling, en moet voldoen aan toegankelijkheidseisen, wat in het onderwijs extra gewicht heeft gezien de diversiteit van de doelgroep.

Bijlage 1: Ethische leidraad voor gebruik AI en data

Mensenrechten en ethiek

Het toetsingskader richt zich vooral op juridische en technische aspecten van AI. Maar het gebruik van AI in het onderwijs brengt niet alleen technische en juridische uitdagingen met zich mee, maar ook ethische verantwoordelijkheden. Gebruikte AI moet veilig zijn en de grondrechten van mensen respecteren. AI kan al snel impact hebben op:

- Recht op onderwijs (art. 14 Handvest): AI mag geen belemmering vormen voor toegang tot of kwaliteit van onderwijs.
- Non-discriminatie (art. 21 Handvest): AI-systemen moeten vrij zijn van bias en discriminatie, bijvoorbeeld bij toelating, beoordeling, of tracking.
- Transparantie en uitlegbaarheid: Leerlingen en ouders moeten begrijpen hoe AI-systemen beslissingen nemen en welke gegevens worden gebruikt.
- Menselijke tussenkomst: Bij hoog-risicotoepassingen (zoals examens of toelating) moet altijd menselijk toezicht mogelijk zijn.

Uit onderzoeken blijkt namelijk dat ‘foute AI’ grote impact kan hebben op de levens van mensen. AI kan grote voordelen bieden, maar het brengt ook risico's mee zoals onterechte profilering, manipulatie, discriminatie en verlies van menselijke controle over beslissingen die mensen direct raken.

De verordening introduceert daarom een risicogebaseerd kader: hoe groter het risico op schade, hoe strenger de eisen. Daarmee wil de wetgever innovatie mogelijk houden, maar wel binnen ethische en juridische grenzen die burgers beschermen.

Voor onderwijsinstellingen is dit relevant omdat zij AI inzetten voor kwetsbare groepen (leerlingen) bij wie fouten of misbruik van AI ingrijpende gevolgen kunnen hebben voor hun ontwikkeling en toekomst. Dit betekent dat scholen niet alleen moeten voldoen aan wettelijke eisen (zoals de AI-Verordening en AVG), maar ook actief moeten werken aan bewustwording en kritisch denken over AI-toepassingen. De inzet van AI hangt daarmee nauw samen met ethiek¹¹. Bij de inzet van AI is het daarom verstandig breder te kijken dan alleen dit toetsingskader, en de totale impact van AI. Kennisnet heeft voor scholen

¹¹ <https://www.kennisnet.nl/artificial-intelligence/ethische-dilemmas-en-de-inzet-van-ai-in-het-onderwijs/>

een 'Handreiking AI'¹² ontwikkelt die scholen helpt bij de inzet van AI, evenals de Waardenwijzer¹³ en Ethiekkompas¹⁴.

Ethische richtlijnen voor leveranciers en scholen

Voor leveranciers en scholen zijn er verschillende ethische afwegingskaders beschikbaar bij het ontwikkelen van educatieve AI en AI-systemen.

Voor ontwikkelaars van AI en AI-systemen zijn de Ethische richtlijnen voor betrouwbare kunstmatige intelligentie¹⁵ beschikbaar. Ook als een AI-toepassing niet onder AI-Verordening valt of lijkt te vallen, verdient het aanbeveling deze richtlijnen te volgen. De ethische richtlijnen voor betrouwbare AI beschrijven de belangrijkste principes om ervoor te zorgen dat AI-systemen wettig, ethisch en robuust zijn. Ze benadrukken menselijke zeggenschap, technische veiligheid, privacy, transparantie, eerlijkheid, maatschappelijk welzijn en verantwoordelijkheid bij de ontwikkeling en implementatie van AI. De richtlijnen zijn opgesteld door de High-Level Expert Group van de Europese Commissie.

Bij gebruik van AI staat menselijke controle centraal. Hierbij worden drie vormen onderscheiden:

1. human-in-the-loop (HITL) beoordeelt een mens elke individuele uitkomst vóórdat die effect heeft.ergelijkbaar met de menselijke tussenkomst uit art. 22 AVG.
2. Bij human-on-the-loop (HOTL) handelt het systeem zelfstandig maar monitort een mens de werking en kan ingrijpen, wat aansluit bij het menselijk toezicht uit art. 14 AI Act.
3. Bij human-in-command (HIC) heeft een mens de strategische regie over de inzet van het systeem als geheel, inclusief de beslissing om het níet te gebruiken.

Dit onderscheid is cruciaal omdat een school kan menen dat "er een mens bij zit" terwijl er in werkelijkheid sprake is van automation bias: iemand klikt op akkoord zonder echt te beoordelen. Dit maakt zichtbaar welk niveau van betrokkenheid een specifiek AI-systeem vereist en of dat in de praktijk ook werkelijk wordt ingevuld. Menselijke tussenkomst is de term van artikel 22 AVG: het is verboden om volledig geautomatiseerd besluiten te nemen die rechtsgevolg hebben of iemand significant raken. Er moet dus een mens zijn die het besluit neemt.

¹² <https://www.kennisnet.nl/artificial-intelligence/handreiking-ai-in-het-onderwijs/>

¹³ <https://www.kennisnet.nl/nieuws/geactualiseerde-waardenwijzer-koers-houden-op-publieke-waarden-bij-digitalisering/>

¹⁴ <https://www.kennisnet.nl/tools/ethiekkompas/>

¹⁵ <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

Voor scholen zijn soortgelijke richtlijnen opgesteld¹⁶. Deze ethische overwegingen zijn intrinsiek waardevol en begeleiden docenten en schoolleiders bij hun beslissingen over het gebruik van AI-tools en algemene AI-modellen in het onderwijs. Het blijft van wezenlijk belang om bij het inzetten van AI voor leren, lesgeven en beoordelen de risico's en context centraal te stellen. De vijf ethische criteria zijn samengevat:

- Menselijke waardigheid (human dignity) – AI moet de intrinsieke waarde van iedere persoon respecteren, inclusief privacy, autonomie en menselijke handelingsvrijheid.
- Billijkheid / eerlijkheid (fairness) – AI moet bijdragen aan gelijke kansen, inclusie, non-discriminatie en een eerlijke verdeling van rechten en verantwoordelijkheden.
- Vertrouwen / betrouwbaarheid (trust/trustworthiness) – AI moet op een eerlijke, transparante, betrouwbare en privacy-respecterende manier worden ingezet, in het belang van leerlingen en goed onderwijs.
- Academische integriteit (academic integrity) – AI moet eerlijk en verantwoord worden gebruikt, zonder misleiding, met correcte bronvermelding en met behoud van geldige toetsing en menselijke oordeelsvorming.
- Gerechtvaardigde keuze (justified choice) – keuzes over AI moeten berusten op kennis, feiten en data, en tot stand komen via transparante, participatieve en uitlegbare besluitvorming.

Verantwoord inzetten van AI in de onderwijspraktijk betekent, zoals door de Europese Commissie benadrukt in de *Guidelines on the ethical use of artificial intelligence* (Europese Commissie, 2026) ook:

- Afstemming op onderwijsdoelen: AI moet leerprocessen verrijken en aansluiten bij de visie van de school.
- Actieve betrokkenheid van stakeholders: Leerlingen, ouders en docenten moeten geïnformeerd en getraind zijn in het gebruik van AI.
- Kritische reflectie: AI mag nooit de menselijke verantwoordelijkheid vervangen, maar moet worden ingezet als ondersteunend instrument.

¹⁶ Updated Guidelines on the ethical use of artificial intelligence and data in teaching and learning for educators (Europese Commissie November 2025, updated March 2026)

Bijlage 2: Checklist algoritmes

Los van de vraag of de AI-toepassing onder de definitie in en van de AI-Verordening valt, brengt het gebruik van algoritmes (de basis van AI) ook risico's met zich mee. Bij AI-systemen die niet onder de AI-Verordening vallen, is het wel raadzaam om te controleren of het algoritme en gebruik van het AI-systeem veilig (genoeg) is.¹⁷

Nr.	Principe	Toelichting	Toelichting voor toetsing
1	Proportionaliteit en 'Do No Harm' (Voorkomen van schade)	Het gebruik van AI mag niet verder gaan dan noodzakelijk is voor een legitiem doel. Er moet vooraf een zorgvuldige risicoanalyse plaatsvinden om mogelijke schade voor individuen of groepen te voorkomen.	→ Is het doel legitiem en proportioneel? → Is er vooraf een risicoanalyse gedaan (DPIA / ethische toets)? → Zijn maatregelen getroffen om schade te voorkomen?
2	Veiligheid en beveiliging	AI-systemen mogen geen onnodige veiligheidsrisico's opleveren en moeten beschermd worden tegen kwetsbaarheden en aanvallen.	→ Zijn veiligheids- en cybersecurityrisico's in kaart gebracht? → Zijn passende beveiligingsmaatregelen genomen?
3	Recht op privacy en gegevensbescherming	Privacy moet in alle fasen van de AI-levenscyclus worden beschermd en bevorderd. Er moet een adequaat juridisch en organisatorisch gegevensbeschermingskader bestaan (AVG/UAVG).	→ Worden beginselen van gegevensminimalisatie, doelbinding en beveiliging toegepast? → Is er een DPIA uitgevoerd? → Wordt naleving van AVG/UAVG geborgd?
4	Meerdere belanghebbenden en adaptief bestuur	Bestuur van AI moet inclusief zijn en de deelname van diverse belanghebbenden mogelijk maken. Internationale normen en nationale soevereiniteit moeten worden gerespecteerd.	→ Zijn relevante stakeholders betrokken (burgers, experts, onderwijsprofessionals, leerlingen, ouders)? → Is governance ingericht op aanpassing bij veranderende risico's?
5	Verantwoordelijkheid en	AI-systemen moeten controleerbaar, auditeerbaar en herleidbaar zijn. Er	→ Is duidelijk wie verantwoordelijk is voor beslissingen en fouten?

¹⁷ UNESCO Recommendation on the Ethics of AI

	aansprakelijkheid	moeten mechanismen bestaan voor toezicht, effectbeoordeling, due diligence en audits om conflicten met mensenrechten en milieuschade te voorkomen.	→ Zijn audit- en toezichtmechanismen ingericht? → Is er documentatie die herleidbaarheid mogelijk maakt?
6	Transparantie en uitlegbaarheid	AI-systemen moeten in een mate die past bij de context transparant en uitlegbaar zijn. Er moet rekening worden gehouden met spanningen tussen uitlegbaarheid en andere beginselen (zoals privacy of veiligheid).	→ Is de werking van het algoritme begrijpelijk uitgelegd (technisch, beleidsmatig, richting betrokkenen)? → Is documentatie zoals model cards of datasheets aanwezig? → Zijn de beperkingen en aannames van het systeem bekend?
7	Menselijke controle en eindverantwoordelijkheid	AI mag menselijke verantwoordelijkheid niet vervangen. Er moet altijd sprake blijven van betekenisvolle menselijke tussenkomst en eindverantwoordelijkheid.	→ Is er voorzien in menselijke toetsing, override of goedkeuring? → Is duidelijk wie de eindbeslissing neemt en aansprakelijk is?
8	Duurzaamheid	AI moet worden beoordeeld op zijn impact op duurzame ontwikkeling, in lijn met de VN-doelstellingen voor duurzame ontwikkeling (SDG's).	→ Is er aandacht voor energieverbruik, milieu-impact en sociale duurzaamheid? → Draagt het systeem bij aan duurzame maatschappelijke doelen?
9	Bewustwording en digitale geletterdheid	Het publiek moet inzicht krijgen in AI en datagebruik via open en toegankelijke educatie, burgerparticipatie, digitale vaardigheden en AI-ethiektraining.	→ Is er beleid om gebruikers, personeel en betrokkenen te informeren en te scholen? → Is er transparante communicatie over werking en risico's van het systeem?
10	Rechtvaardigheid en non-discriminatie	AI moet sociale rechtvaardigheid, eerlijkheid en non-discriminatie bevorderen. De voordelen van AI moeten voor iedereen toegankelijk zijn.	→ Zijn bias- en discriminatierisico's systematisch getoetst en gemonitord?

			→ Worden fairness-principes toegepast in ontwerp en evaluatie? → Zijn kwetsbare groepen meegenomen in de impactbeoordeling?
--	--	--	---

Bijlage 3: FAQ en aanbevelingen

1. Kan een aanbieder het beoogde doelbewust smal formuleren om hoog-risico-classificatie te vermijden?

Nee. De AI-Verordening sluit dit op meerdere manieren uit.

Ten eerste telt niet alleen het opgegeven beoogde doel, maar ook de functie van het AI-systeem, modaliteiten van gebruik en redelijkerwijs te voorzien misbruik mee (art. 9 lid 2 sub b en overweging 65). Een aanbieder die weet dat scholen een AI-module zullen gaan gebruiken voor overgangsbeslissingen, kan zich niet verschuilen achter de omschrijving "administratieve ondersteuning." Het beoogde doel moet bovendien consistent zijn in alle communicatie: gebruiksinstructies, reclame- en verkoopmateriaal én technische documentatie (art. 3 lid 12).

Ten tweede kijken nationale markttoezichtautoriteiten naar de feitelijke technische werking van het AI-systeem. Als die overeenkomt met een hoog-risicocategorie uit Bijlage III, prevaleert dat boven de opgegeven beschrijving. In dat geval voldoet de leverancier niet aan de verplichtingen voor AI-systemen met een hoog risico.

Ten derde werkt er een vangnet van de ketenverantwoordelijkheid: de school is als gebruiksverantwoordelijke verplicht het AI-systeem in te zetten volgens de gebruiksinstructies. Als de school dat niet doet, en het AI-systeem (dat niet als hoog risico was geclassificeerd) wél voor een hoog-risico-AI-toepassing inzet, dan wordt de school zelf gezien als aanbieder (art. 25) met alle bijbehorende verplichtingen.

Ten slotte is het opzettelijk verkeerd etiketteren, en het verstrekken van onjuiste informatie, onvolledige of misleidende informatie, een zelfstandige overtreding in de AI-Verordening (boetecategorie 15 miljoen euro/3% wereldwijze jaarlijkse omzet respectievelijk 7,5 miljoen euro/1% wereldwijze jaarlijkse omzet).

2. Als docenten AI-tools of generatieve AI zoals ChatGPT, Copilot of Mistral zelf aanpassen voor onderwijstoepassingen, kan de school daardoor zelf als aanbieder van dat AI-systeem worden aangemerkt en is dat een probleem?

Ja, onder omstandigheden. Het enkele gebruik van een bestaande AI-tool zoals ChatGPT, Copilot of Mistral door docenten maakt de school meestal nog geen aanbieder, maar een gebruiksverantwoordelijke. De AI-Verordening kent wel een specifieke bepaling voor situaties waarin een bestaand AI-systeem wordt aangepast voor een nieuw doel: degene die die aanpassing doet, kan worden aangemerkt als aanbieder in de zin van de

verordening (art. 25). Omdat de docent in dienst is bij de school, vallen deze activiteiten onder verantwoordelijkheid van de school.

Dit risico is concreet relevant bij het gebruik van generatieve AI (zoals Copilot, Claude of ChatGPT). Wanneer een docent of school een AI-systeem via een systeem-prompt, fine-tuning of geavanceerde inrichting of integratie inricht voor een specifieke toepassing, bijvoorbeeld het beoordelen van huiswerk of toetsen of opstellen van overgangsadviezen, en die toepassing valt onder een hoog-risicocategorie uit Bijlage III, dan kan de school als aanbieder worden aangemerkt. Dat geldt ook bij vibecoding waarbij er op een bestaand AI-platform wordt geprogrammeerd, doorgaans is dat platform niet ingericht op onderwijsspecifieke situaties en risico's.

De school is dan zelf verantwoordelijk voor de conformiteitsbeoordeling, technische documentatie, registratie in de EU-databank en alle overige verplichtingen die normaal bij de oorspronkelijke leverancier liggen. Deze gevolgen voor de school zijn, los van eventuele handhaving en boetes, groot evenals de risico's voor de gezondheid, veiligheid en fundamentele rechten van leerlingen en medewerkers. Uiteraard zal de school niet aan deze verplichtingen kunnen gaan voldoen.

De ketenverantwoordelijkheid geldt ook als de school geen commercieel oogmerk heeft. De verordening kijkt naar de functie en het gebruik, niet naar de juridische of commerciële status van de inzetter.

Let op: het gaat er om dat de school méér doet dan normaal gebruik of lichte configuratie van een AI-systeem, AI-model of generatieve AI (zoals ChatGPT). Zie voor meer toelichting de uitleg in de paragraaf "Verschuiving van aanbiedersverantwoordelijkheid" in het hoofdstuk over hoog-risico-AI.

3. Is een AI-systeem altijd het gehele softwarepakket, of kan het een afzonderlijk onderdeel zijn?

Een AI-systeem kan binnen de AI-Verordening zowel een afgebakend onderdeel van een groter ict- of softwarepakket zijn als het gehele geïntegreerde systeem of softwarepakket (ongeacht of het geïntegreerd is of losgekoppeld functioneert). Dit volgt ook expliciet uit de AI-Verordening zelf: AI-systemen die onderdeel zijn van grootschalige IT-systemen vallen onder de AI-Verordening, ook als het bredere systeem al vóór 2027 op de markt was gebracht. De kwalificatie en verplichtingen kleven aan het juridisch relevante AI-systeem. Dat kan een afgebakende AI-component zijn, maar ook het geïntegreerde geheel, afhankelijk van wat precies op de markt wordt gebracht of in gebruik wordt genomen en met welk beoogd gebruik.

Voor AI-systemen met hoog risico betekent dit het volgende. Als een groter softwarepakket meerdere functies bevat en slechts één afgebakend AI-subsysteem is bedoeld voor bijvoorbeeld beoordeling, toelating of examinering (hoog risico), dan is in beginsel alleen dat AI-subsysteem hoog-risico. De rest van het pakket wordt niet door nabijheid automatisch meegetrokken. De classificatie hangt immers af van de functie, het beoogde gebruik en de modaliteiten van gebruik van het AI-systeem zelf.

Maar als het geïntegreerde geheel als één AI-systeem op de markt wordt gebracht voor een hoog-risicofunctie, of het gebruik van het gehele pakket is verbonden met de AI-module, dan is dat gehele systeem hoog-risico. De scheidslijn loopt dus niet langs de technische grens tussen "hele applicatie" en "module," maar bij de relevante vraag: wát wordt er precies in gebruik genomen, en met welk beoogd doel? Een geïntegreerd systeem kan daarmee als geheel hoog-risico zijn als het beoogde gebruik dat rechtvaardigt.

Voorbeeld: bij de beoordeling van een LAS met AI-functionaliteit is de eerste vraag niet "is dit pakket hoog-risico?" maar "welk AI-systeem of welke AI-component zit hierin, en wat is het beoogde gebruik daarvan?".

Bronvermelding en verantwoording

De totstandkoming van dit Toetsingskader AI voor het funderend onderwijs is mogelijk dankzij subsidie van het ministerie van OCW. De verschillende conceptversies zijn opgesteld in samenwerking en nauwe afstemming met Kennisnet. In verschillende stadia van de ontwikkeling van het toetsingskader heeft overleg en afstemming plaatsgevonden met Edu-V, PO-Raad, VO-raad, NOLAI, NPULS en SURF, alsmede brancheorganisaties van leveranciers MEVW, VDOD en een aantal leveranciers die de (ontwikkeling van) hun producten met SIVON deelden. SIVON is alle partijen erkentelijk voor hun medewerking.

Dit Toetsingskader AI is mede gebaseerd op de inzichten uit de volgende publicaties:

- VERORDENING (EU) 2024/1689 VAN HET EUROPEES PARLEMENT EN DE RAAD van 13 juni 2024 tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie en tot wijziging van de Verordeningen (EG) nr. 300/2008, (EU) nr. 167/2013, (EU) nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 en (EU) 2019/2144, en de Richtlijnen 2014/90/EU, (EU) 2016/797 en (EU) 2020/1828 (*verordening artificiële intelligentie*)
- *Updated Guidelines on the ethical use of artificial intelligence and data in teaching and learning for educators* (Europese Commissie November 2025, updated March 5th 2026)
- *Making informed choices on digital education content: EU guidelines for teachers and educators* (Europese Commissie, 2025)
- *Richtsnoeren van de Commissie betreffende verboden artificiële-intelligentiepraktijken zoals vastgesteld bij Verordening (EU) 2024/1689 (AI Verordening)*, Europese Commissie, C2025 5052 final.
- *Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)*, (EC July 2025)
- *Assessing high-risk artificial intelligence - fundamental rights risks* (European Union Agency for Fundamental Rights, 2025)
- *AI Verordening Beslshulp* (MinBZK, maart 2026) <https://AI-Verordening-beslshulp.apps.digilab.network/>
- *Tijdljn AI Verordening Algoritmekader* <https://minbzk.github.io/Algoritmekader/AI-Verordening/tijdljn-AI-Verordening/>
- *European Commission - Questions and answers Artificial Intelligence* (EC August 2024)
- *Gids AI-Verordening* (versie 1.1 september 2025, Ministerie van EZ)
- *Cloud Sovereignty Framework* (ver 1.2.1 Oct 2025)

- *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01), (EDPB 2018)*
- *Navigating the AI Act – FAQ* <https://digital-strategy.ec.europa.eu/en/faqs/navigating-ai-act>
- *The AI Board Workshop on High-Risk Classification – Education and vocational training – Annex III Point 3 (EC June 2025)*
- *AI Act digital strategy – AI ACT FAQ (EC jan 2026),* <https://digital-strategy.ec.europa.eu/policies/regulatory-framework-ai>
- *Shaping Europe’s digital future (EC 2026),* <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- *AI Act Service desk FAQ (Oct 2025),* <https://ai-act-service-desk.ec.europa.eu/nl/faq>
- *Navigating the AI Act, FAQ, (EC 2026),* <https://digital-strategy.ec.europa.eu/en/faqs/navigating-ai-act#ecl-inpage-high-risk-ai-systems>