

Stel jezelf deze vijf vragen voordat je AI inzet voor nakijkwerk

Als er één taak is die je als docent liever minder zou doen, dan is het waarschijnlijk nakijken. Hoe fijn zou het zijn als we dat kunnen automatiseren?



Over de auteurs

Eva de Schipper is wetenschappelijk onderzoeker bij CitoLab (Stichting Cito) en producteigenaar van CheckMate. Wouter Jansen is leraar NaSk op het Corbulo College (Lucas Onderwijs) in Voorburg. Als teacher in residence is hij ook betrokken bij NOLAI.

Gezien het grote tekort aan leerkrachten¹, de hoge werkdruk² en de dalende internationale rankings van ons onderwijs³, is het automatiseren van nakijkwerk allesbehalve een gek idee. AI-technologie lijkt hierin een uitkomst te kunnen bieden. Vooral werkstukken waarbij leerlingen zelf een antwoord formuleren, zoals essays, verslagen en toetsen met open vragen, geven waardevol inzicht in wat leerlingen hebben geleerd. Helaas kost juist het nakijken van deze werkstukken veel tijd, is het foutgevoelig en wordt het vaak als saai en eentonig ervaren.^{4,5,6}

Al tientallen jaren wordt er onderzoek gedaan naar het automatisch nakijken van open antwoorden.^{7,8,9} De recente doorbraak van generatieve AI heeft de ontwikkeling en introductie van nakijktechnologieën op de markt in een stroomversnelling gebracht. CitoLab onderzoekt nakijkondersteuning met het prototype CheckMate.¹⁰ Learnbeat¹¹ en Kwizl hebben nakijkondersteuning ingericht in hun platformen en er zijn initiatieven als Nakijken.ai¹² of Toets-Tester¹³. Tot slot wordt ook generatieve AI 'stand alone' gebruikt als nakijktechnologie. In de meeste genoemde gevallen blijft nog wel een antwoordmodel of rubric nodig.

The screenshot shows the CheckMate interface. On the left, a task card for 'Salamander' (Salamander) provides background information and a question. The question asks for an abiotic factor from the text. The text describes the kamsalamander (amphibian) and its habitat. A photo of a salamander is included. On the right, the 'ANTWOORDMODEL' (Answer Model) is 'water'. Below this, 'LEERLINGANTWOORDEN' (Student Answers) are listed. The first answer is 'water', which is marked correct with a green dot. Below it, several other answers are listed, each with a red dot indicating they are incorrect. The answers are: 'leeft vooral in de buurt van bossen en heggen', 'Hij leeft vooral in de buurt van bossen en heggen.', 'bossen en heggen', 'leven in buurt van bossen en heggen', and 'nestgelegenheden'.

Salamander

De kamsalamander is zeer zeldzaam in Nederland. Hij leeft vooral in de buurt van bossen en heggen. De voortplanting vindt plaats in water. Een vrouwtje legt dan zo'n 200 eieren. De larven eten watervlooien.

Voor het voortbestaan zijn voor de kamsalamander zowel biotische als abiotische factoren belangrijk.

Noteer de abiotische factor uit de tekst die voor de kamsalamander belangrijk is.

ANTWOORDMODEL 1

water

LEERLINGANTWOORDEN SUGGESTIES 1

5x **water** X ✓ 1

Water. ✓

leeft vooral in de buurt van bossen en heggen •

Hij leeft vooral in de buurt van bossen en heggen. •

bossen en heggen •

leven in buurt van bossen en heggen •

nestgelegenheden •

Screenshot van nakijkondersteuning CheckMate, een prototype van CitoLab.

De ontwikkelingen op het gebied van AI bieden kansen voor het onderwijs, maar er zijn ook redenen om voorzichtig te zijn. CitoLab en NOLAI hebben daarom samen deze whitepaper opgesteld over nakijken met AI. Hiermee geven wij leerkrachten en schoolleiders vijf kritische vragen mee. Deze vragen zijn vooral bedoeld om vanuit de eigen specifieke schoolsituatie aan te zetten tot nadenken en overwegen, voordat tot de inzet van AI in het nakijkproces wordt overgegaan.

Vraag 1

Welke impact heeft geautomatiseerd nakijken op je pedagogiek en didactiek?

De allerbelangrijkste vraag voor iedere onderwijsvernieuwing, zou je kunnen zeggen.¹⁴ Wetenschappelijk gezien is hier nog bar weinig over bekend. Toch spelen hier tal van vraagstukken. Is de kwaliteit van de nakijktechnologie goed genoeg om leerlingantwoorden zelfstandig te beoordelen? Hoe kritisch blijven leerkrachten nog nakijken als hun AI-nakijkhulp 'redelijk' goed nakijkt?¹⁵ Gaan leerlingen hun antwoorden 'tweaken' richting de AI-nakijkmodellen?¹⁶ Krijgen we 'eenheidsworst' in de toetsen, waarbij we vooral het type vragen in toetsen gaan opnemen dat zich door AI laat nakijken? Wie heeft hier het afbreukrisico als er fouten worden gemaakt, de AI of de leerkracht? Hoe goed kan een AI overweg met leerlingen die hun antwoord net even anders formuleren, bijvoorbeeld omdat Nederlands hun tweede taal is?¹⁷ Hoe hou je als docent 'gevoel' bij het gemaakte werk van iedere leerling? Een hele toets nakijken is immers iets anders dan alleen het eindcijfer zien. Dashboards kunnen helpen om toch inzicht in het proces van de leerling te behouden, maar hoe moet zo'n dashboard er dan uit zien?

"Op welke manier houd je inzicht in het leerproces van je leerlingen?"

Vraag 2

Doet de technologie wat ik nodig heb?

Het is belangrijk om na te gaan of de nakijktechnologie nakijkt op een manier die jou helpt. Ga daarvoor na welke functionaliteiten jij nodig hebt en of dat verandert in verschillende contexten, zoals summatief en formatief toetsen, of toetsen op paper of digitaal. Hoe automatisch wil je dat het nakijken gaat? Alles geheel automatisch nagekeken, een score-advies per antwoord of alleen vergelijkbare antwoorden groeperen, zodat je makkelijker kan nakijken? Hoe transparant moet de tool zijn over waarom hij doet wat hij doet? En op welke manier houd je inzicht in het leerproces van je leerlingen nu je minder (of misschien zelfs niets) van de gemaakte toets zelf ziet?

Naast de nakijkfunctionaliteiten zelf waar je op kunt letten, is bijvoorbeeld ook de ondersteuning van verschillende soorten toetsvragen belangrijk of hoe makkelijk je toetsen kan invoeren. Maar ook het verkrijgen van uitgebreide rapportages, feedback voor de leerling of de mogelijkheid om de resultaten makkelijk naar andere systemen over te zetten zijn van belang. Daarnaast moet de nakijktechnologie natuurlijk gebruiksvriendelijk zijn en is er zelfs dan misschien wel training nodig voor de docent. Het is namelijk

een vereiste om goed AI-geletterd te zijn als je met hoog-risico AI-toepassingen werkt. Vraag je dus af: in hoeverre moet je de AI snappen om met AI-nakijktechnologie te werken?

Vraag 3

Mag nakijken met AI eigenlijk wel?

Europese wetgeving rondom AI-technologie, de zogenoemde 'AI-Act', is vanaf 2024 van kracht.¹⁸ Over nakijken schrijft deze wetgeving¹⁹ het volgende:

AI-systemen met een hoog risico op grond van artikel 6, lid 2, zijn de vermelde AI-systemen op de volgende gebieden (...): 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;

Nakijken door AI – een vorm van evaluatie – blijkt dus in de hoog-risicocategorie te vallen. Dat betekent dat AI-nakijkhulp alleen toegepast mag worden als er aan de vereisten van hoog-risico-AI wordt voldaan. Dat vraagt om o.a. een systeem voor risicomanagement en kwaliteitsbeheer, een CE-keurmerk, een FRIA (Fundamental Rights Impact Assessment), transparantie en uitlegbaarheid van de werking van AI, menselijk toezicht, gegevensbeheer van de data en post market monitoring.²⁰ Geen sinecure dus. Deze situatie wordt anders als de nakijktechnologie alleen voorbereidende werkzaamheden verricht en het uiteindelijke oordeel bij de docent blijft. Het is echter nog niet duidelijk wat precies onder 'voorbereidend' valt. Wees daarom kritisch op de software en de partij die dit aanbiedt als je nakijktechnologie op je school implementeert.

In alle gevallen blijft het belangrijk dat je als docent moet kunnen afwijken van een AI-oordeel zonder dat direct te hoeven verantwoorden naar het systeem of de schoolleiding. Maar in hoeverre ervaar je die ruimte nu echt als er straks nakijktechnologie beschikbaar is?

Vraag 4

Help je als school mee om de nakijkmodellen te trainen?

Bij sommige AI-nakijktools die de markt op komen, worden scholen actief gevraagd om mee te ontwikkelen. Vaak met de belofte dat zij dan de komende jaren gratis gebruik mogen maken van de tool. Voor die ontwikkeling zijn dan 'geanonimiseerde nagekeken toetsen' nodig, tezamen met de originele toets en het nakijkmodel. Daarmee kunnen dan de AI-nakijktools getraind worden.

De vraag is echter van wie het intellectueel eigendom is van de antwoorden op de toets (die door de leerling zijn geformuleerd) en van het correctiewerk (dat door de leerkracht is verricht). Bij dat laatste hoogstwaarschijnlijk de school, bij het eerste hangt dat wellicht af van het soort antwoord. Een simpel antwoord op een wiskundevraag kent wellicht geen intellectueel eigendom, maar een essay voor Nederlands misschien wel. Het doel van de antwoorden en het correctiewerk was een beoordeling van het werk, wat voor beide partijen vooraf bekend was. In hoeverre mogen deze data aan derden gegeven worden voor een ander doel – en los van of het juridisch mag, wat vinden we ervan? Daarbij is ook nog de vraag wie de organisatorische zaken rondom deze data gaat regelen (denk aan anonimiseren en zodanig ordenen dat deze verwerkt kan worden), want ook dit is een tijdrovende bezigheid. Tot slot maken veel scholen ook gebruik van standaard methodetoetsen en ook de uitgeverijen daarvan hebben auteursrecht op hun werk.

“Zelfs bij geanonimiseerde verwerking bestaat altijd de kans op herleiding.”

In het verlengde hiervan ligt natuurlijk de vraag hoe de beveiliging en de privacy van de verstrekte gegevens wordt gewaarborgd. Zelfs bij geanonimiseerde verwerking bestaat altijd de kans op herleiding. En het is ook pijnlijk voor een school als je geanonimiseerde toetsen per ongeluk op straat belanden. Zijn er verwerkingsovereenkomsten met de partijen die de nakijktools organiseren en hebben zij hun informatiebeveiliging op orde?

Vraag 5

In hoeverre is AI ook een ethische oplossing?

De tot nu toe geziene nakijktechnologie is vaak op basis van machine learning of large language modellen. Rondom large language modellen, zoals ChatGPT, Copilot of Gemini, zijn de nodige ethische vraagstukken. Er lopen rechtszaken rondom inbreuk op auteursrecht. Het opschonen van de modellen, om de ergste fouten en ongepastheden eruit te halen, wordt vaak in lagelonenlanden gedaan door mensen die daardoor veelvuldig worden blootgesteld aan hele ingrijpende content. Het energieverbruik en de milieubelasting is groot. Het feit dat de modellen grootschalig gebruikt worden, ontslaat ons als onderwijs – en dus publieke partij – niet van de verantwoordelijkheid om te onderzoeken of de technologie in lijn is met onze waarden. De overheid heeft hiervoor al behulpzame richtlijnen opgesteld (in dit geval wel specifiek over het gebruik van generatieve AI).²¹

Hoe nu verder?

Wij roepen bestuurders, schoolleiders en leerkrachten op om kritisch na te denken over deze vijf vragen voordat ze zelf aan de slag gaan met nakijktechnologieën. De verleiding om snel in te stappen is groot en de wens om de hoeveelheid nakijkwerk te reduceren is heel begrijpelijk. AI-technologie lijkt daar op het eerste gezicht nu al bij te kunnen helpen.

“De technologie werkt op zichzelf behoorlijk goed, maar we zijn er nog niet.”

Wij hopen je met deze vijf kritische vragen aan het denken te zetten. De technologie werkt op zichzelf behoorlijk goed, maar we zijn er nog niet. Er is nog veel nodig om te komen tot een duurzame, ethisch en wettelijk verantwoorde, gebruikersvriendelijke en betrouwbare nakijkhulp voor leerkrachten.

Daarom nodigen we iedereen uit om hier samen aan te werken. Bijvoorbeeld door deel te nemen aan wetenschappelijk onderzoek. Begin gerust met veilige, weldoordachte experimenten op je eigen school, maar denk kritisch na en laat je begeleiden door experts. Laten we samen zorgen dat deze nakijktechnologie tot bloei komt.

Tijdens de Nationale Onderwijs Tentoonstelling (NOT) 2025 presenteerden Eva en Wouter het co-creatieproject ‘Open vragen nakijken met hulp van AI’.



Bronnen

1. [Aanpak lerarentekort | Werken in het onderwijs | Rijksoverheid.nl](#)
2. [Werkdruk in het Onderwijs - 1 - Monitor arbeid](#)
3. [Een blik op de Nederlandse positie in internationale onderwijsrankings](#)
4. Sachin Kumar, Soumen Chakrabarti, and Shourya Roy. (2017). Earth mover's distance pooling over siamese lstms for automatic short answer grading. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, pages 2046-2052.
5. [Burrows, S., Gurevych, I. & Stein, B. \(2015\). The Eras and Trends of Automatic Short Answer Grading. International Journal of Artificial Intelligence in Education 25, 60–117.](#)
6. Toetsing, beoordeling en feedback: omschrijving en toelichting van veelgebruikte begrippen door SLO. Raadpleegbaar op: slo.nl/thema/meer/curriculum-toetsing/
7. Valenti, S., Neri, F., Cucciarelli, A. (2003). An overview of current research on automated essay grading. Journal of Information Technology Education: Research, 2(1), 319-330.
8. [Burrows, S., Gurevych, I., Stein, B. \(2015\). The eras and trends of automatic short answer grading. International Journal of Artificial Intelligence in Education, 25\(1\), 60-117.](#)
9. [Bai, X., Stede, M. \(2023\). A Survey of Current Machine Learning Approaches to Student Free-Text Evaluation for Intelligent Tutoring. International Journal of Artificial Intelligence in Education 33, 992–1030.](#)
10. <https://cito.nl/onderzoek-innovatie/portfolio/innovatieve-prototypes/de-waarde-van-data/checkmate/>
11. [De nakijkassistent - Learnbeat Support](#)
12. <https://nakijken.ai/>
13. <https://toetstester.nl/>
14. [Díaz, B., & Nussbaum, M. \(2024\). Artificial intelligence for teaching and learning in schools: The need for pedagogical intelligence. Computers & Education, 105071](#)
15. [Dell'Acqua, Fabrizio \(2022\). Falling asleep at the wheel: Human/AI collaboration in a field experiment on HR recruiters.](#)
16. [Filighera, A., Ochs, S., Steuer, T. et al. \(2024\). Cheating Automatic Short Answer Grading with the Adversarial Usage of Adjectives and Adverbs. International Journal of Artificial Intelligence in Education 34, 616–646.](#)
17. [Idowu, J.A. \(2024\). Debiasing Education Algorithms. International Journal of Artificial Intelligence in Education 34, 1510–1540.](#)
18. <https://www.kennisnet.nl/artificial-intelligence/voldoen-aan-de-ai-verordening/>
19. <https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX%3A32024R1689>
20. <http://www.ictrecht.nl/hubs/Marketing%20bestanden/AI/ICTRecht%20Cheat%20Sheet%20AI%20Act.pdf>
21. [Overheidsbrede handreiking voor de verantwoorde inzet van generatieve AI | Rapport | Rijksoverheid.nl](#)