

onderdeel van Innovatieagenda 2026-2029



Voorwoord

De wereld van onderwijs en toetsing is voortdurend in beweging. Wat we meten, hoe we dat doen en wat we daarmee zichtbaar maken, verandert mee. Dat vraagt om scherpe keuzes en nieuwe perspectieven. Want als de context verandert, kan onze aanpak niet achterblijven.

Waar kansengelijkheid vroeger vooral betekende dat iedere leerling toegang had tot hetzelfde onderwijs en via vaste routes kon doorstromen, verschuift de aandacht nu naar de vraag of die routes ook recht doen aan de verschillen tussen leerlingen. In een klas vol uiteenlopende achtergronden, talen en talenten werkt één standaardaanpak niet meer voor iedereen. Het gaat om eerlijke kansen: ruimte maken voor verschil, zodat iedere leerling kan groeien op zijn of haar eigen manier.

Tegelijkertijd verandert het onderwijs snel door technologische ontwikkelingen, met AI als belangrijke versneller. Dat beïnvloedt hoe leerlingen leren en hoe leraren werken. Het biedt kansen om werkdruk te verlagen en leren beter te ondersteunen, maar vraagt ook om zorgvuldige keuzes. Want het onderwijs is geen ‘experimenteertuin’. Vernieuwing moet doordacht zijn, met oog voor kwaliteit en verantwoordelijkheid.

In die dynamiek zoeken wij als divisie CitoLab, onderdeel van Stichting Cito, continu naar verdieping en verbinding. Om die zoektocht te sturen, bundelen we onze onderzoeks- en innovatieprojecten in een werkprogramma. Daarin werken we intensief samen met universiteiten en andere kennisinstellingen om onze inzichten te verrijken. Ook werken we samen met leraren in de praktijk. Door die samenwerkingen halen we ‘diepere’ en praktische kennis in huis en laten we ons kritisch bevragen.

De afgelopen jaren ontwikkelden we innovatieve meetinstrumenten en pasten die toe in onderzoek met duizenden leerlingen. We lieten zien wat er mogelijk is. De volgende stap is om die innovaties ook echt in de dagelijkse onderwijspraktijk te brengen. Want pas wanneer ze worden gebruikt door een leraar, in een leerling-volgsysteem of een examen, maken ze het verschil voor leerlingen. We werken de

komende tijd daarom toe naar bredere toepassing en een grotere groep gebruikers. Niet om groter te worden, maar om meer impact te maken.

De kern blijft dezelfde vraag: hoe maken we leren zichtbaar op een manier die recht doet aan verschillen tussen leerlingen? De thema's uit de innovatieagenda bieden deze mogelijkheid om antwoord te vinden op deze vraag. Toetsing is daarbij een middel om eerlijkere kansen en beter inzicht in de ontwikkeling te bieden.

Dit werkprogramma laat zien hoe wij vanuit CitoLab bijdragen aan de Cito Innovatieagenda 2026-2029, samen met partners en het onderwijsveld. Gedreven door één overtuiging: iedereen verdient een eerlijke kans om zich te ontwikkelen.

Want alleen door leren zichtbaar te maken, kunnen we echt het verschil maken.

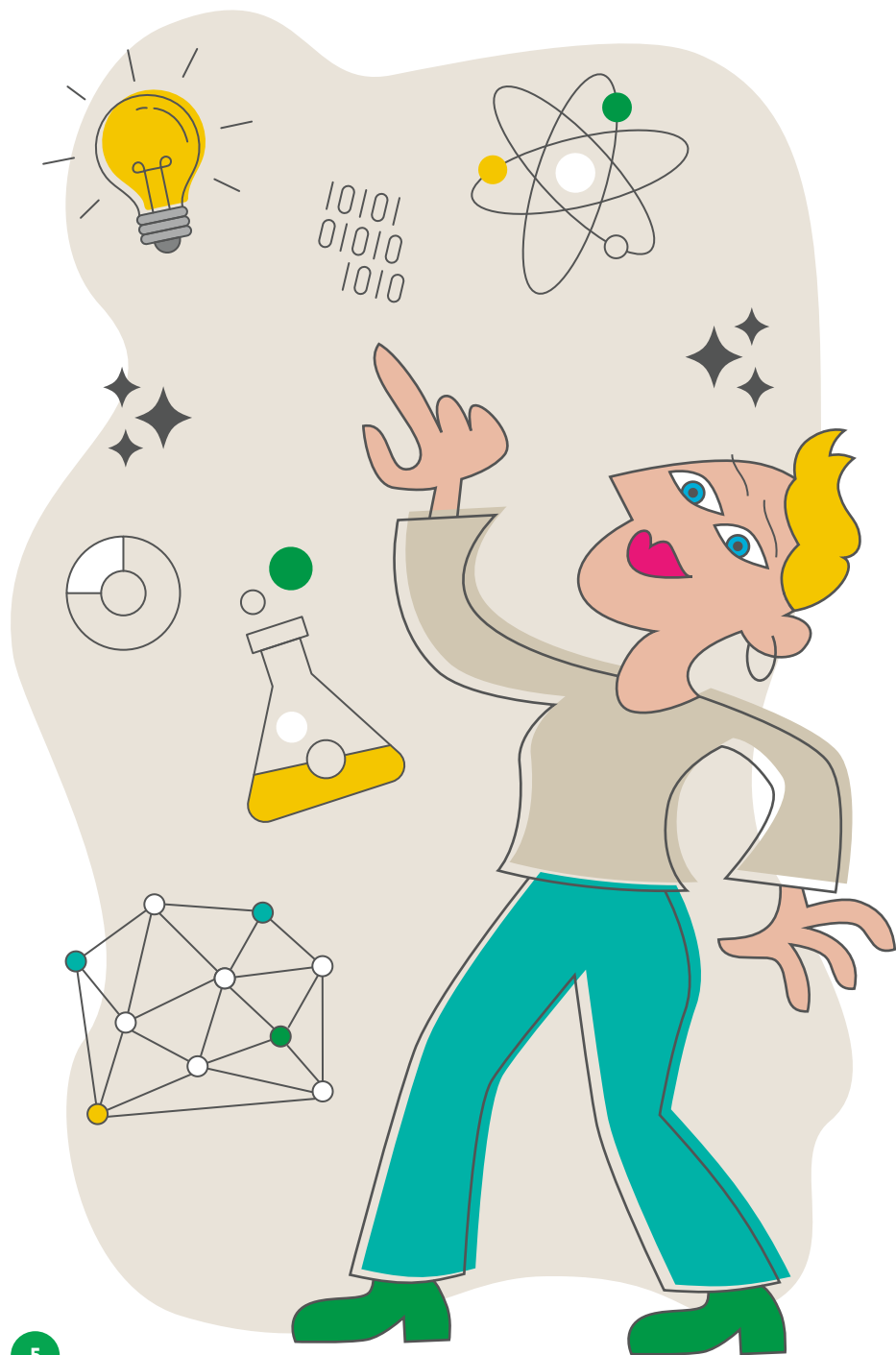
Remco Feskens

Directeur CitoLab | Stichting Cito



Inhoud

	Voorwoord	2
	Missie voor CitoLab	6
	Krachtig samen, voor en met het onderwijs	8
	Van thema's naar activiteiten	10
1	Innovatiethema Groeien als jezelf	12
2	Innovatiethema Verrijkende metingen	18
3	Innovatiethema Buiten gewone vakken	28
4	Innovatiethema Balans tussen mens en machine	34
	Tot slot	46
	Colofon	47



Missie voor CitoLab

CitoLab draagt met praktijkgericht en ontwerpgericht onderzoek bij aan de missie van Cito: leren zichtbaar maken. Dat doen we niet alleen door nieuwe kennis en tools te ontwikkelen, maar ook door die actief te delen met het onderwijsveld en collega's binnen Cito. In dit hoofdstuk vertellen Romy Noordhof (hoofd team prototypes) en Emmelien van der Scheer (hoofd team onderwijskundig onderzoek) hoe het programma uit de innovatieagenda hieraan bijdraagt. Ze gaan in op de kracht van programmagericht werken, de ontwikkelingen en kansen die nu spelen in het onderwijsveld, en de impact die CitoLab de komende jaren wil maken.

Samenhang als sprong voorwaarts

“We gaan nu het negende jaar in van het werken met werkprogramma's en aan alles merken we dat dit programma echt kan bijdragen aan onze missie”, vertelt Romy. “Dat komt vooral doordat we de stap maken van een set projecten naar een samenhangend programma.”

Volgens Emmelien maakt juist die samenhang het verschil. “We bouwen voort op de ervaring die we de afgelopen jaren hebben opgedaan. Door onze groei als divisie weten we beter waar onze kracht ligt, welke expertise we in huis hebben en waar we nog verder in moeten leren.”

Die ontwikkeling zorgt voor meer slagkracht dan voorheen. Ook helpt het dat de Cito-brede missie – leren zichtbaar maken – steeds duidelijker richting geeft aan de activiteiten van CitoLab.

Een maatschappelijke behoefte: basis versus breed

Het onderwijs en de maatschappij zijn volop in beweging. Enerzijds groeit de behoefte om breder naar de ontwikkeling van leerlingen te kijken dan alleen naar basisvaardigheden. Anderzijds staan juist die basisvaardigheden in Nederland onder druk. Ook de ‘klassieke’ vorm van toetsing wordt steeds vaker ter discussie gesteld. “Met CitoLab willen we inspelen op dat krachtenveld”, zegt Emmelien. “We onder-

zoeken hoe je op een valide en betrouwbare manier zicht kunt krijgen op de brede ontwikkeling van leerlingen, én hoe je anders kunt kijken naar het volgen van basisvaardigheden.”



Dat vraagt om nieuwe oplossingen en nieuwe manieren van meten. Op korte termijn ontwikkelt CitoLab prototypes en praktische toepassingen. Op de middellange termijn willen we onderbouwde antwoorden bieden op vragen als: hoe je vaardigheden goed in beeld brengt en verantwoord binnen scholen inzet.

De nieuwe standaard: het is niet voor de toets

“Gestandaardiseerde toetsen bieden veel zekerheid en vertrouwen”, zegt Romy. “Er is weinig ruimte voor beoordelingsfouten en één goed antwoord geeft duidelijkheid.” Tegelijkertijd past die duidelijkheid niet altijd bij vaardigheden die leerlingen later in de praktijk nodig hebben.

“Ik zou het mooi vinden als we over vier jaar met hetzelfde vertrouwen kunnen kijken naar meer authentieke vormen van toetsen”, vervolgt Romy. “Bij een spellingstoets is het logisch dat je spelling meet. Maar hoe mooi zou het zijn als we met hetzelfde vertrouwen ook vaardigheden als probleemoplossend vermogen of creativiteit kunnen meten?”

CitoLab onderzoekt daarom hoe bestaande kennis over toetsen en beoordelen gecombineerd kan worden met nieuwe technologieën, zoals simulaties, data en AI. Niet als doel op zich, maar als ondersteuning om rijkere en meer open leerling-antwoorden goed, eerlijk en betrouwbaar op te kunnen halen en te beoordelen. De ambitie is groter dan alleen nieuwe toetsvormen ontwikkelen. Uiteindelijk moet de scheiding tussen leren en toetsen steeds kleiner worden. Toetsing wordt dan niet meer ervaren als iets los van het leren, maar als een vanzelfsprekend onderdeel van het leerproces.

Door vanuit verschillende perspectieven naar deze vraagstukken te kijken en praktische oplossingen te ontwikkelen, wil CitoLab richting geven aan complexe, vakoverstijgende uitdagingen in het onderwijs.

Krachtig samen, voor en met het onderwijs

De impact van onze activiteiten ontstaat niet alleen door een gecoördineerde set van projecten en een heldere missie. Minstens zo belangrijk zijn de samenwerkingen met scholen, onderzoekers en andere partners. Samen zorgen zij ervoor dat onderzoek en innovaties aansluiten op de praktijk van vandaag en bijdragen aan het onderwijs van morgen. Door praktijkvragen te verbinden met wetenschappelijke kennis ontstaan oplossingen die relevant, bruikbaar en onderbouwd zijn. Zo kunnen innovaties sneller landen in het onderwijs en beter aansluiten bij de behoeften van leerlingen, leraren en scholen.

Vreemde ogen, zo vreemd nog niet

De samenwerking met externe partners is daarin onmisbaar. Onderzoekers brengen diepgaande kennis mee, terwijl scholen inzicht geven in de dagelijkse onderwijspraktijk. In combinatie met de kennis van CitoLab over leren zichtbaar maken, leidt dat tot oplossingen die zowel theoretisch sterk als praktisch toepasbaar zijn.

CitoLab werkt daarom samen met verschillende universiteiten, waaronder de universiteiten van Twente, Nijmegen, Utrecht en sinds kort ook Amsterdam. Met de Radboud Universiteit wordt bijvoorbeeld intensief samengewerkt op het gebied van leesvaardigheid. Een thema dat binnen toetsing en onderwijs sterk in ontwikkeling is. Die samenwerking biedt toegang tot onderzoekers die zich al jarenlang verdiepen in vragen als: wat is leesvaardigheid precies, en hoe meet je die op een goede en betrouwbare manier? Met de universiteit van Amsterdam werken we samen aan het vernieuwen van de data-analyse die achter leerlingvolgysteemtoetsen ligt. Dat is belangrijk om nieuwe vormen van toetsing ook met statistische zekerheid bij scholen te kunnen implementeren; we willen dat de informatie die je aan leraren geeft over een leerling, informatie is waar ze op kunnen vertrouwen én dus mee aan de slag kunnen.

De school als leerplek voor innovatie

Naast wetenschappelijke kennis zijn ook de vragen en ervaringen uit scholen van grote waarde. Bij het ontwikkelen van prototypes, het testen van innovaties en het verzamelen van data is CitoLab regelmatig aanwezig op scholen: in de klas, in de lerarenkamer en in gesprekken met leraren en leerlingen. Juist daar ontstaan vaak inzichten die onderzoek verder brengen.

Een goed voorbeeld daarvan is CitoLab on Tour. Tijdens een bezoek van twee dagen aan een vmbo-school testten leerlingen en leraren verschillende prototypes. Hoewel de innovatieve oplossingen veel enthousiasme opriepen, kwamen de belangrijkste vragen juist vanuit de dagelijkse praktijk. “Jullie oplossing voor het nakijken van open antwoorden is veelbelovend”, vertelde een leraar. “Maar hoe kijk ik straks voegwoorden na?”

Dat soort gesprekken maakt duidelijk waar de echte behoefte ligt. Schoolbezoeken zijn daarom niet alleen waardevol voor het testen van innovaties, maar ook een leerplek voor CitoLab zelf: een plek waar zichtbaar wordt welke vragen, uitdagingen en behoeften er spelen in het onderwijs.

Van pilots naar blijvende impact

Soms vraagt innovatie juist om klein te beginnen: testen, leren en verbeteren in de praktijk. Tegelijkertijd wil CitoLab ook groter denken als het gaat om impact. Hoe zorgen we ervoor dat kennis, inzichten en tools niet beperkt blijven tot pilots of onderzoeken, maar daadwerkelijk hun weg vinden naar het onderwijs?

De afgelopen jaren heeft CitoLab al stappen gezet van kleinschalige pilots naar onderzoeken met duizenden leerlingen. De ambitie voor de komende jaren is om innovaties verder door te ontwikkelen en te implementeren in leerlingvolgsystemen, examinering en de dagelijkse onderwijspraktijk.

De projecten binnen het werkprogramma zijn daarop ingericht: samen ontwikkeld met scholen, onderzoekers en experts, met als doel om oplossingen te creëren die niet alleen vernieuwend zijn, maar ook duurzaam toepasbaar in het onderwijs.

Van thema's naar activiteiten

De komende vier jaar geven we onze projecten vorm vanuit vier thema's uit de innovatieagenda:

- 1 **Groeien als jezelf**
- 2 **Verrijkende metingen**
- 3 **Buiten gewone vakken**
- 4 **Balans tussen mens en machine**

Met ons onderzoek en onze innovaties werken we aan het waarmaken van de beloften van deze thema's. Dat doen we vanuit onze expertise in onderwijskundig meten en altijd in samenwerking met anderen. Samen zoeken we naar oplossingen voor de vraagstukken van vandaag én morgen.

Uit ervaring weten we dat de meest succesvolle projecten het gedachtegoed van meerdere thema's combineren. Dat is niet verrassend: in de praktijk zijn activiteiten, projecten en doelen onlosmakelijk met elkaar verbonden.

Op de volgende pagina's lees je meer over de vier thema's en de projecten die daarbij horen. Tussendoor vind je interviews met CitoLab-collega's die tijdens het schrijven van dit werkprogramma al volop aan de slag gingen.

Veel leesplezier!



1 Innovatiethema | Groeien als jezelf

Leerlingen moeten kunnen leren en groeien op een manier die bij hen past. Iedereen brengt zijn eigen ervaringen, talenten en uitdagingen mee naar school. De groei die ze doormaken, verloopt zelden volgens een strak, lineair patroon.

Persoonlijke leerpaden worden daarom steeds belangrijker dan uniforme routes richting succes. Zeker in een meer diverse samenleving. Dit vraagt om een verschuiving in denken: van relatieve normen – die vooral meten wie er boven of onder het gemiddelde zit – naar inzicht bieden in waar een leerling staat en hoe die de volgende stap kan zetten.

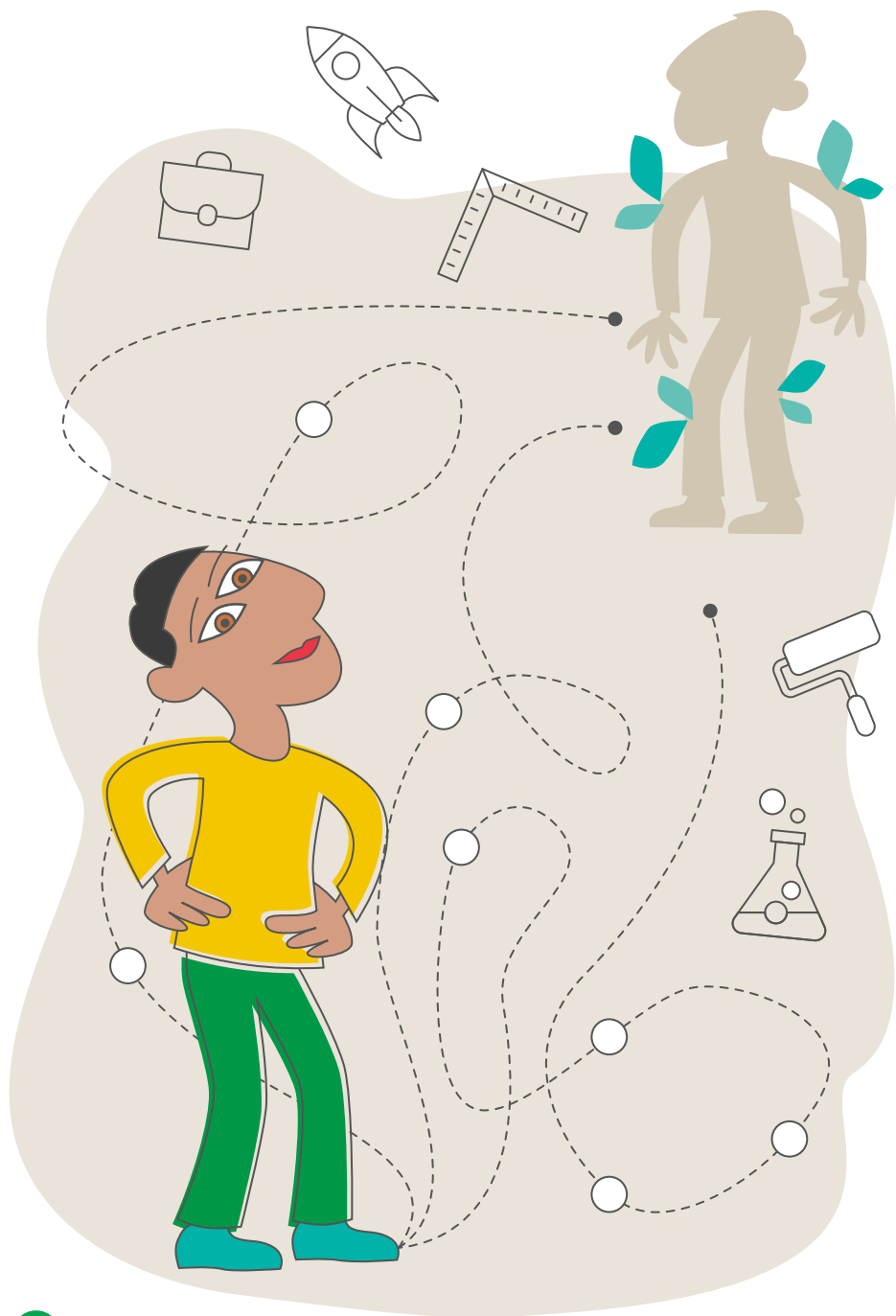
We laten hoge verwachtingen niet los. We maken ze persoonlijk en concreet, zodat leerkrachten én leerlingen weten welke haalbare volgende stappen er zijn. Daarmee bieden we eerlijke kansen waarin alle leerlingen zich kunnen ontwikkelen.

Het gaat om toegankelijke leeromgevingen waar diversiteit niet als een probleem wordt gezien, maar als een kracht. Zo wordt leren betekenisvol en motiverend, terwijl prestatiedruk vermindert.

Onderwijsvraagstukken die in dit thema centraal staan:

- Hoe kunnen we toetsvormen ontwerpen die recht doen aan de veelzijdigheid en diversiteit van leerprocessen?
- Hoe kunnen we leerlingen stimuleren hun ontwikkeling zichtbaar te maken, zonder de prestatiedruk te vergroten?
- Hoe kunnen we toetsen en beoordelen herontwerpen, zodat ze meer gericht zijn op groei en ontwikkeling dan op het simpelweg vaststellen van een prestatieniveau.

Onze belofte: Wij ontwikkelen producten die leren zichtbaar maken, zodat je met vertrouwen de volgende stap kunt zetten.



Spelenderwijs werken aan zelfregulatie



Aranka
Bijl



Ingrid
Williams

In gesprek met Aranka Bijl (onderwijskundig onderzoeker) en Ingrid Williams (R&D-specialist), die samen werken aan Rekenbaas: een prototype dat inzicht geeft in je eigen leerproces.

Binnen CitoLab werken Aranka en Ingrid samen aan Rekenbaas. Een prototype met een spelelement waarbij het vermogen om jezelf in te schatten en je leerstrategie bij te sturen centraal staat. In dit interview vertellen zij over het ontstaan van het project, de keuzes die ze maakten en de volgende stappen.

Van idee naar papieren prototype

De aanleiding voor Rekenbaas ligt in een bredere ontwikkeling in het onderwijs, schetst Aranka. “Zelfregulerend leren krijgt steeds meer aandacht, omdat deze verzameling van cognitieve vaardigheden belangrijk is voor schoolsucces. Toch ligt de focus vaak nog sterk op inhoud: wat kan een leerling wel of niet? Inzicht krijgen in hoe leerlingen hun leren sturen, is lastiger.” Met Rekenbaas wilden zij dat perspectief kantelen. “Het gaat niet om hoe goed je rekt, maar om hoe goed je jezelf inschat. Kun je voorspellen of een opgave je gaat lukken? En wat doe je als dat anders uitpakt dan verwacht?”

Rekenen als context, niet als doel

Hoewel het spel nu draait om rekensommen, is rekenen vooral de context. “Je moet ergens opdrachten aan ophangen,” zegt Aranka. “Maar in de kern gaat het om zelfinschatting. Daarom hebben we het zo opgezet dat het in principe ook toepasbaar is op andere vakken, zoals taal of geschiedenis.” De naam Rekenbaas verwijst naar dat idee: de baas zijn over je eigen leren. Rekenbaas is eerst ontwikkeld voor het primair onderwijs. “Maar we hebben het ook getest met scholieren uit het voortgezet onderwijs. De sommen waren te makkelijk, maar ze vonden het spel wel leuk.” Dat bevestigt het idee dat het concept schaalbaar is.

“Als je de inhoud en moeilijkheid aanpast, kan het voor veel doelgroepen werken,” vult Ingrid aan. “We hebben zelfs al een VO-school die interesse heeft getoond.”

Van spel naar betekenisvolle feedback

De recente focus lag op rapportage: wat krijgt een leerling eigenlijk terug na het spelen? “We wilden feedback geven die gaat over zelfinschatting,” legt Aranka uit, “en leerlingen stimuleren om het spel vaker te spelen.”

Daarbij schuilt ook een uitdaging. “Leerlingen zijn gewend om feedback te koppelen aan goed of fout,” zegt Ingrid. “Terwijl wij juist willen dat ze begrijpen: dit gaat niet over hoe goed je bent in rekenen, maar over hoe je jezelf inschat.” Ook voor leraren wordt gewerkt aan rapportages. Het doel is niet alleen inzicht, maar ook handelingsperspectief. “We onderzoeken of we met profielen kunnen werken,” vertelt Ingrid. “Bijvoorbeeld voor leerlingen die zichzelf structureel overschatten. Dan kan een leerkracht gerichte ondersteuning bieden.”

Vooruitblik: testen, uitbreiden en valideren

Dit jaar staan verdere uitbreiding en validatie centraal. De rapportages worden verfijnd, er komt mogelijk een extra domein bij en Aranka en Ingrid hopen het prototype op grotere schaal te testen in de klas. “Dat zou ons helpen om te zien of de aanpassingen echt bijdragen aan het inschattingsvermogen van leerlingen,” zegt Aranka. Ook zien zij kansen richting uitgeverijen en andere onderwijs-ontwikkelaars. Rekenbaas laat zien hoe zelfregulerend leren op een toegankelijke en speelse manier inzichtelijk kan worden gemaakt. Ingrid: “Dat zie je nog nauwelijks terug in bestaande tools. Juist dat maakt dit project zo interessant.”

Wat maakt dit project bijzonder?

Voor Aranka zit het vooral in het spelontwerp. “Het is leuk om steeds te blijven nadenken over hoe je het spelelement sterk houdt. En tegelijk bieden we iets nieuws: een concreet antwoord op de vraag hoe je zelfregulatie zichtbaar kunt maken.” Ingrid sluit zich daarbij aan. “Die combinatie van spel en zelfregulerend leren zie je nog niet vaak. Dat we dit samen hebben kunnen ontwikkelen, maakt dit project voor mij echt bijzonder.”

Waarderen wat groeit |

Nieuw project

Iedere leerling ontwikkelt zich op een eigen manier in zijn of haar schoolloopbaan. Gestandaardiseerde toetsen helpen scholen om die ontwikkeling betrouwbaar in beeld te krijgen, maar staan soms ook ver van de praktijk. Wanneer we de ontwikkeling van leerlingen rijker kunnen meten, zal dit de leerkracht helpen en de leerling nog meer recht doen. Uiteraard moet dit wel op een betrouwbare en valide manier gebeuren, moet het waardevol en haalbaar zijn voor leerkrachten. In dit project onderzoeken we hoe we op een valide, betrouwbare en waardevolle manier leerlingen nog zorgvuldiger kunnen volgen in de tijd. Zo kun je de dingen waar we waarde aan hechten in het onderwijs nog meer waarderen!

Autonoom naar eindniveau |

Nieuw project

Leerlingen op Agora-scholen hebben in de onderbouw veel inspraak in het bepalen van hun leerinhoud en een bijpassende toetsingsvorm. Maar in de overgang naar de bovenbouw wordt steeds meer inhoud bepaald en mede daardoor grijpen docenten ook sneller terug op de traditionele toetsvormen. Hoe kun je leerlingen stimuleren om authentieke toetsvormen te kiezen,

waarbij je als docent kunt garanderen dat ze toewerken naar het vereiste eindniveau? Deze vraag staat centraal in dit project: we ontwikkelen samen met een docent authentieke toetsvormen, waarbij leerlingen kunnen laten zien dat ze hun vaardigheden op het gewenste niveau beheersen.

Snelle signalering, diepgaand gesprek | *Nieuw project*

Grip krijgen op de sociaal-emotionele ontwikkeling van leerlingen is lastig: gedrag varieert per dag en gebeurt in een context. De resultaten van metingen hangen hierdoor sterk af van wie de meting invult en wanneer. Hierdoor is het interpreteren van één meting weinig zeggend, en het interpreteren van een trend ingewikkeld. Daarbij vraagt het invullen van deze metingen veel tijd van leerkrachten en leerlingen. In dit project onderzoeken we een nieuw doel: het signaleren als het niet goed gaat, in plaats van de focus op het volgen van de ontwikkeling. We ontwikkelen een betrouwbare, periodieke screeningsaanpak die vroegtijdig signaleert of het mogelijk niet goed gaat. Van daaruit kan een leerkracht dieper met de leerling in gesprek.

Verder komen met hulp |

Nieuw project

Meestal leren we met een toets iets over de vaardigheid van de leerling, zoals welk onderwerp nog wel wat aandacht kan gebruiken. Door dynamisch te toetsen kun je iets zeggen over welke hulp een leerling nodig heeft om verder te komen. De leerling kan bijvoorbeeld genoeg hebben aan een aanmoediging om nog eens goed naar de vraag te kijken, of juist een specifieke hint in de richting van een oplossing nodig hebben. In dit project onderzoeken we wat ervoor nodig is om de hulpbehoefte van leerlingen in kaart te brengen door het maken van een toets. Welke hulp hebben zij nodig om verder te komen?

Rekenbaas | *Doorontwikkelingsproject*

Rekenbaas is een prototype dat leerlingen ondersteunt bij het ontwikkelen van zelfregulerend leren. Leerlingen geven bij rekenopgaven aan of zij denken dat hun antwoord klopt en hoe zeker zij daarvan zijn. Door directe feedback en een spelelement vergroot Rekenbaas het zelfinzicht en stimuleert het doelgericht leren. In 2026 zetten we in op twee vervolgstappen. We breiden de aanpak van zelfregulatie uit naar een ander domein dan rekenen. Daarnaast valideren we samen met scholen het onderliggende model van Rekenbaas en



onderzoeken we hoe zelfregulerend leren breder en effectiever kan worden gestimuleerd.

Toetsen als brug |

Doorontwikkelingsproject

In brede brugklassen is toetsen een uitdaging. Hoe ontwerp je een toets die past bij meerdere niveaus en toch goed onderscheid maakt? Nu gebeurt dit vaak met grensscores, bijvoorbeeld bij een 7,5 ga je naar vwo, daaronder naar havo. Andere scholen gebruiken geen cijfers voor het maken van de beslissing. Wat is een betrouwbare manier om leerlingen aan het best passende schoolniveau toe te wijzen? In 2025 zochten we uit hoe scholen kiezen voor brede brugklassen en wat de impact hiervan is op hoe je toetst. In 2026 ontwikkelen we een praktische handleiding die we valideren in samenwerking met scholen.

2 Innovatiethema | Verrijkende metingen

In een snel veranderde wereld verdienen onze toetsen een upgrade. Denk aan verrijkende opdrachten waarin je de vaardigheden zo goed als meteen kunt aflezen, waarvan je de validiteit kunt doorvoelen, die een doorkijkje geven naar de vaardigheid buiten schooltijd en die vaardigheden integreren als dat van meerwaarde is.

En niet zonder reden: in de veranderende wereld met veranderende banen in een ongewisse toekomst, komt er nadruk op andere vaardigheden. Wat als we het volgen en zichtbaar maken van deze vaardigheden aanvliegen met onze kennis van nu, zonder onze historie van meten de boventoon te laten voeren?

We gebruiken slimme technologieën om vernieuwende meetinstrumenten te ontwikkelen voor vaardigheden die we eerder niet (afdoende) maten. Verrijkende metingen is een thema dat draait om het maken van keuzes die leren en toetsen echt verbinden met de maatschappij.

Onderwijsvraagstukken die in dit thema centraal staan:

- Hoe kunnen we beoordelingspraktijken ontwikkelen die recht doen aan complexe, geïntegreerde vaardigheden zoals probleemoplossing, creativiteit en samenwerken of productieve vaardigheden zoals spreken?
- Hoe kunnen we technologie inzetten om real-time, bruikbare feedback te geven aan leerlingen en docenten zonder verlies van didactische betekenis?
- Hoe kunnen we formatieve en summatieve metingen zo integreren dat toetsen het leren ondersteunen, zonder het ongewenst te sturen?

Onze belofte: We ontwikkelen vernieuwende meetinstrumenten die leren en toetsen verrijken en verbinden aan de maatschappij.



“Bij Navigeren naar Leesbegrip onderzoeken we vraagvormen die dieper leesbegrip zichtbaar maken en specifieke obstakels diagnosticeren.”

Aranka Bijl, projectleider Navigeren naar Leesbegrip

Een digitale gesprekspartner



Marica
Balk



Renée
Venema

In gesprek met Marica Balk (R&D-specialist) en Renée Venema (onderzoeks-stagiaire) over het project dat AI inzet om gesprekken te voeren.

‘In gesprek met AI’ draait om de vraag hoe AI kan bijdragen aan het inzichtelijk maken van gespreksvaardigheid. Gespreksvaardigheden in kaart brengen kent namelijk veel haken en ogen. Zo kan de onderlinge relatie tussen leerlingen of tussen de leerling en leraar hun gesprek sterk beïnvloeden, en dat is niet wat je inzichtelijk wil maken. Ook vraagt het voeren van 1 op 1 gesprekken veel tijd van docenten. Het doel is om een betrouwbare en natuurlijke gesprekspartner te ontwikkelen. Marica en Renée vertellen hoe het project ontstond, wat ze tot nu toe ontdekten en wat de volgende stap is.

AI als nieuwe component

“Eigenlijk begon het idee jaren geleden al,” vertelt Marica. “We hadden toen al projecten met taal- en spraaktechnologie, bijvoorbeeld bij het prototype ‘Beeldverhaal’. Maar echte interactieve gespreksvaardigheid zat daar nauwelijks in. We merkten dat er veel vragen van logopedisten en NT2-docenten kwamen: kunnen we dit gebruiken? Maar de technologie was nog niet ver genoeg. We merken nog steeds dat de praktijkvraag groot is. Toen ontstond de vraag: kunnen we nu, met AI, wel een tool ontwikkelen die in gesprek gaat met leerlingen?” Renée vult aan: “Toen Marica mij benaderde, dacht ik meteen: dit past goed bij mijn NT2-achtergrond. Mijn eerste stap was literatuuronderzoek naar de vraag: hoe wordt gespreksvaardigheid nu gemeten? Het antwoord: nauwelijks, en nooit op schaal. Dat maakt het arbeidsintensief en duur. AI biedt juist de kans om een gesprekspartner beschikbaar te maken die altijd klaarstaat. Het beoordelen kan later; eerst wilden we weten of AI überhaupt een gesprek kan voeren.”

Van idee naar data

Het team besloot zich te richten op NT2-leerders. “We hebben acht verschillende situaties ontwikkeld, face-to-face en telefonisch, die we eerst met leerlingen zelf uitvoerden en opnamen. Met deze opnames konden we testen of AI de spraak goed kan verstaan en hoe we prompts moeten opstellen om een gesprek op gang te houden.”

Renée: “Tijdens de dataverzameling heb ik samen met andere NT2-docenten gesprekken gevoerd met NT2-sprekers. Ik probeerde te testen of AI een passende vervolgvraag kan stellen, kan samenvatten wat de deelnemer zegt of bevestigen wat er is gezegd. Soms werkt het goed, soms niet, en daar probeer ik steeds de prompts op aan te passen. Soms moet de AI ook gewoon zeggen: ‘oké’ of ‘dat begrijp ik’, en dat blijkt lastig.”

Van onderzoek naar prototype

Parallel aan het onderzoek ontwikkelden zij een eerste prototype. “Het prototype is nog basaal,” vertelt Marica. “De stem klinkt robotachtig, de responstijd is vertraagd. Maar het laat ons wel zien dat spraak verwerkt kan worden en dat AI inhoudelijk kan reageren. Zo krijgen we inzicht in wat werkt en waar de knelpunten zitten: snelheid, intonatie, interactie.”

Vervolg vragen en gebruikerservaring

Het project roept ook ontwerp vragen op over de vorm en beleving van de AI-gesprekspartner. “Wil je een avatar? Moet de stem mannelijk, vrouwelijk of afwisselend zijn? Hoe ervaren gebruikers het gesprek?” vertelt Renée. “Sommige deelnemers vinden het prettiger met AI te praten dan met een mens, anderen vinden het juist spannend of onnatuurlijk. Dat varieert heel erg per persoon, en dat is interessant voor vervolgonderzoek.”

Vooruitblik: testen en doorontwikkelen

De komende fase staat in het teken van doorontwikkeling en gebruikerstests. Het doel is een gebruiksvriendelijk werkend prototype. Daarbij ligt de focus eerst op de inzet van privacybestendige speech-to-speech-technologie. Daarna volgen de verdere ontwikkeling van de AI-gesprekspartner en AI-ondersteuning bij de beoordeling.

Van stem naar vaardigheid |

Nieuw project

Spreekvaardigheid is essentieel voor succesvolle deelname aan de maatschappij en speelt een belangrijke rol bij bijvoorbeeld het leren lezen. In het onderwijs is er weinig aandacht voor het inzichtelijk maken van spreekvaardigheid, omdat het beoordelen veel tijd kost en snel subjectief wordt. In dit project onderzoeken we hoe spreekvaardigheid op een valide manier gemeten kan worden en welke rol AI daarbij kan spelen. We onderzoeken of AI kan helpen om transparant en automatisch te beoordelen, wat kan leiden tot eerlijkere beoordelingen en gerichte, automatische feedback voor leerlingen.

Navigeren naar leesbegrip |

Doorontwikkelingsproject

De manier waarop we leesvaardigheid meten, krijgt steeds meer aandacht. Experts benadrukken dat goed lezen gaat over meer dan woorden begrijpen. Leerlingen moeten een mentaal beeld kunnen vormen van wat ze lezen, ook wel een situatiemodel genoemd. De traditionele benadering 'lees-de-tekst-en-maak-de-vragen' moedigt leerlingen waarschijnlijk niet aan zo'n situatiemodel op te bouwen, andere vraagvormen en taakinrichtingen lijken

daarvoor beter geschikt. In 2025 werkten we aan een gezamenlijke visie op lezen, interviewden we experts om de uitdagingen rond het meten van leesvaardigheid verder te verkennen, en ontwierpen we taken die leerlingen ondersteunen in het opbouwen van een situatiemodel en daarmee hogere-ordevaardigheden zichtbaar maken. Dit bouwen we in 2026 uit door leestaken te onderzoeken die een duidelijk leesdoel geven. Daarnaast onderzoeken we hoe docenten diagnostische informatie kunnen bieden met gerichte leestaken.

In gesprek met AI |

Doorontwikkelingsproject

Gesprekken voeren is een vaardigheid die in de praktijk overal terugkomt en altijd cruciaal is gebleven voor het prettig kunnen meedraaien in de maatschappij. Het beoordelen hiervan is echter zeer tijdsintensief en een objectieve afname staat, zeker bij interactieve gesprekken, onder druk. AI biedt kansen om deze toetsing te verbeteren en minder tijdsintensief te maken. In 2025 zijn we gestart met het onderzoeken én experimenteren in hoeverre AI van toegevoegde waarde kan zijn in de toetsing en het oefenen van mondelinge gesprekssituaties. In 2026 onderzoeken we of AI niet alleen kan reageren op wat je zegt, maar ook kan reageren op hoe je iets zegt.

Hoe kan AI bijdragen aan het zichtbaar maken van spreekvaardigheid?



In gesprek met Marian Marchal (psychometrisch onderzoeker) over het beoordelen van spreekvaardigheid met AI.

Spreekvaardigheid is een belangrijk onderdeel van het beheersen van een (vreemde) taal en een voorwaarde voor leren lezen en schrijven. Het inzichtelijk maken van spreekvaardigheid gebeurt weinig in het onderwijs,

omdat zowel het meten als het beoordelen ervan veel tijd kost. In het onderzoek ‘Van Stem naar Vaardigheid’ zoeken we met behulp van AI naar authentieke toetsvormen om dit te verbeteren. Als eerste stap onderzoekt Marian of AI kan helpen bij het beoordelen van spreekvaardigheid.

Taal, technologie en toekomstperspectief

Voor Marian is het project een mooie combinatie van haar achtergrond en actuele ontwikkelingen. “Ik heb een achtergrond in taal, dus ik vind het superleuk dat ik met taalonderzoek bezig kan zijn, en dan ook nog in combinatie met AI.” Ze ziet bovendien duidelijke kansen voor de toekomst. “Het belangrijkste is dat we de kwaliteit van beoordelingen kunnen verbeteren. Hiermee kan spreekvaardigheid mogelijk een belangrijke plaats krijgen in het onderwijs. Op langere termijn kan dit onderzoek ook bijdragen aan gerichte, automatische feedback voor leerlingen, bijvoorbeeld in oefensituaties ter voorbereiding op examens. Denk aan leerlingen die hun mondeling Engels oefenen met een tool die inzicht geeft in hun sterke- en ontwikkelpunten. Wat ik er zo leuk aan vind, is dat we nu in de beginfase zitten van iets dat uiteindelijk heel waardevol kan worden voor het onderwijs.”

Betrouwbaar én uitlegbaar: een cruciale randvoorwaarde

“Met AI onderzoeken we of we beoordelaars kunnen ondersteunen bij een consistentere en objectievere beoordeling,” vertelt Marian. Het beoordelen van spreekvaardigheid kost namelijk veel tijd, en twee beoordelaars kunnen tot verschillende oordelen komen. In het project kijkt het team of verschillende aspecten van spreekvaardigheid, zoals grammatica en vloeiendheid, automatisch

beoordeeld kunnen worden. Daarbij staat één voorwaarde centraal: de beoordeling moet uitlegbaar zijn. “We willen begrijpen waarom een systeem tot een bepaalde beoordeling komt,” legt Marian uit. “Alleen dan kunnen we iets zeggen over de validiteit en de bruikbaarheid ervan in de praktijk.”

Van Engelstalige naar Nederlandstalige dataset

We werkten met een Engelstalige spraakdataset van mensen die Engels leren. Dit jaar maken we de stap naar het Nederlands en gebruiken we een dataset van kandidaten die Nederlands als tweede taal leren. We onderzoeken of we met kenmerken van spraak, zoals snelheid en uitspraak, automatisch een voorspelling kunnen doen van het taalniveau. Uiteindelijk blijft de beslissing over het taalniveau in handen van de beoordelaar; de voorspelling kan ondersteunen.

Eenvoud als kracht: verrassend goede resultaten

Wat Marian vooral aanspreekt, is dat het team met een relatief eenvoudige aanpak al verrassend goede resultaten behaalde. Ze kozen bewust voor een uitlegbaar beoordelingsmodel, gebaseerd op het bestaande beoordelingsmodel. Vanuit dat beoordelingsmodel onderzochten ze welke onderdelen gekwantificeerd en automatisch bepaald konden worden, en of dat voldoende was om te voorspellen welke score een kandidaat op een toets zou behalen. “Dat bleek eigenlijk al heel goed te werken,” vertelt Marian. In plaats van een complexe AI-aanpak met een ‘black box’ – waarbij patronen worden gevonden zonder dat duidelijk is waarom iemand bijvoorbeeld goed of minder goed presteert in een tweede taal – koos het team voor een valide en begrijpelijk model. Juist die combinatie maakt het onderzoek veelbelovend. Tegelijkertijd blijft de techniek achter het model complex en vraagt het om een uitleg die ook voor niet-specialisten goed te volgen is.

Meer verdieping?

Het onderzoek van Marian en haar collega’s is in 2025 gepubliceerd in een wetenschappelijke paper. Daarin worden de gekozen aanpak en technische onderbouwing uitgebreider beschreven. De publicatie is vooral interessant voor lezers die zich verder willen verdiepen in de methodiek achter AI-ondersteunde beoordeling van spreekvaardigheid. Lees de paper hier.

Authentiek burgerschap meten | *Nieuw promotieonderzoek*

Dit promotieonderzoek maakt deel uit van het landelijke peilingsonderzoek burgerschapsonderwijs in het primair, voortgezet en speciaal onderwijs. We meten kennis, houdingen en vaardigheden rond burgerschap. Een nieuw element is een prototype waarin leerlingen werken aan een authentiek maatschappelijk dilemma, waardoor hun beslisproces en afwegingen zichtbaar worden. Het prototype wordt in 2026 in meerdere onderwijssoorten getest en op basis hiervan verfijnd. In 2027 volgt de grootschalige data-verzameling, waarmee we inzicht willen krijgen in de effectiviteit van onderwijs-aanpakken en hulpmiddelen voor burgerschapsvorming en aanbevelingen kunnen doen voor beleid en praktijk.

Meten van diep tekstbegrip | *Vervolg promotieonderzoek*

De afgelopen jaren hebben we onderzocht hoe diepgaand leesbegrip kan worden gemeten en hoe dit verschilt tussen tekstgenres. Lezen is een complex proces dat diverse cognitieve vaardigheden vereist. Het is dan ook niet verrassend dat lezers regelmatig met hun gedachten afdwalen. De consensus is dat zulke gedachedwaling samenhangt met een

lager leesbegrip. Er is echter ook een hypothese die stelt dat gedachedwaling tijdens het lezen van verhalende teksten kan bijdragen aan het opbouwen van het mentale model, en daarmee het begrip kan versterken. In 2026 onderzoeken we deze hypothese.

Waarderen van praktische vaardigheden in doorstroom po-vo | *Vervolg promotieonderzoek*

Vanaf 2024 doen we met de Universiteit Twente onderzoek naar het toetsen van praktische vaardigheden in de overgang van primair naar voortgezet onderwijs. We willen scherper krijgen wat in deze context relevante praktische vaardigheden zijn, hoe we die inzichtelijk kunnen maken en hoe die kennis een betekenisvolle rol kan spelen in de schooladvisering naar het voortgezet onderwijs. De afgelopen periode hebben we een grootschalige literatuurstudie uitgevoerd om een beeld te kunnen presenteren van de huidige internationale kennis over toetsing van praktische vaardigheden in het primair en voortgezet onderwijs. Zo kunnen we in 2026 specifieker onderzoek doen naar relevante praktische vaardigheden in de aanloop naar het ontwikkelen van een prototype.

Navigeren naar leesbegrip



In gesprek met Aranka Bijl (onderwijskundig onderzoeker) over diep tekstbegrip en nieuwe manieren van toetsen

Leesvaardigheid staat al jaren in de belangstelling. Internationale onderzoeken laten zien dat Nederlandse leerlingen steeds meer moeite hebben met lezen. Volgens Aranka heeft dit te maken met verschillende factoren, zoals leesmotivatie, leesonderwijs en mogelijk ook de

toetsing van leesbegrip. Bij CitoLab onderzoeken we hoe we de toetsing van leesvaardigheid kunnen verdiepen. We doen hiervoor onderzoek naar het inzichtelijk maken van diep leesbegrip. Diep tekstbegrip is echter lastig te meten. Het is bijvoorbeeld moeilijk om daar vragen voor te maken met één duidelijk goed antwoord.

Van oppervlakkig naar diep tekstbegrip

Nederlandse leerlingen scoren relatief goed op vragen waarvan het antwoord letterlijk in de tekst staat. Maar bij diep tekstbegrip – zoals evalueren, vergelijken, kritisch nadenken en inferenties maken – wordt het lastiger. “Juist deze vaardigheden zijn moeilijk te meten. Daardoor krijgen ze minder aandacht, terwijl dit wel het niveau van leesbegrip is dat we leerlingen willen laten ontwikkelen.”

Het situatiemodel centraal

Het project ‘Navigeren naar leesbegrip’ richt zich op het zogenoemde situatie-model: het mentale beeld dat een lezer tijdens het lezen opbouwt. “Het gaat om begrijpen waar een tekst écht over gaat: welke concepten erin voorkomen, hoe die samenhangen en wat impliciet wordt bedoeld.” Dat mentale beeld wordt beïnvloed door voorkennis en eerdere ervaringen. Ook inferenties spelen een belangrijke rol: lezers vullen zelf informatie aan die niet letterlijk in de tekst staat.

Niet alleen meten, maar ook diagnosticeren

Bij diep tekstbegrip kan een leerling op verschillende punten vastlopen. Dat kan te maken hebben met woordenschat, wereldkennis of inzicht in tekststructuren. Daarom willen we niet alleen beter meten, maar ook beter diagnosticeren: welke onderdelen van het leesproces beheerst een leerling al en waar ontstaan de knelpunten?

Van verkenning naar prototypes

Op basis van interviews met experts en literatuuronderzoek verkennen we verschillende oplossingsrichtingen. Een daarvan is het gebruik van concept maps en graphic organizers, waarmee leerlingen laten zien hoe zij verbanden leggen tussen concepten. Daarnaast ontwikkelen we prototypes met nieuwe leestaken die hogere orde denken zichtbaar maken. “We onderzoeken vraagvormen die verder gaan dan traditionele meerkeuzevragen.”

Lezen met een doel

Een belangrijk uitgangspunt is authentiek lezen. In scenario gebaseerde opdrachten combineren en gebruiken leerlingen informatie uit meerdere teksten om een taak uit te voeren. “Je leest in het dagelijks leven altijd met een doel,” zegt Aranka. “Je leest omdat je iets wilt weten of doen.” Door opdrachten in een betekenisvolle context aan te bieden, wordt beter zichtbaar hoe leerlingen informatie verwerken en betekenis geven aan teksten.

Samen leren en ontwikkelen

De belangstelling voor het project is groot. Zowel collega's binnen Cito als andere onderwijsorganisaties reageren enthousiast. “Er is duidelijk behoefte aan nieuwe manieren om leesvaardigheid te toetsen.”

Wat moet het project opleveren?

Aranka hoopt dat het project uitgroeit tot een inspirerende showcase met voorbeelden van hoe leesvaardigheid ook getoetst kan worden. Met aandacht voor diep tekstbegrip, diagnostische informatie en mogelijk zelfs voor formatieve feedback. Het project richt zich op de bovenbouw van het primair onderwijs, waar leerlingen de stap maken van leren lezen naar lezen om te leren. “Juist dan worden verschillen in woordenschat en kennis zichtbaar en wordt betekenis cruciaal.”

Vooruitblik

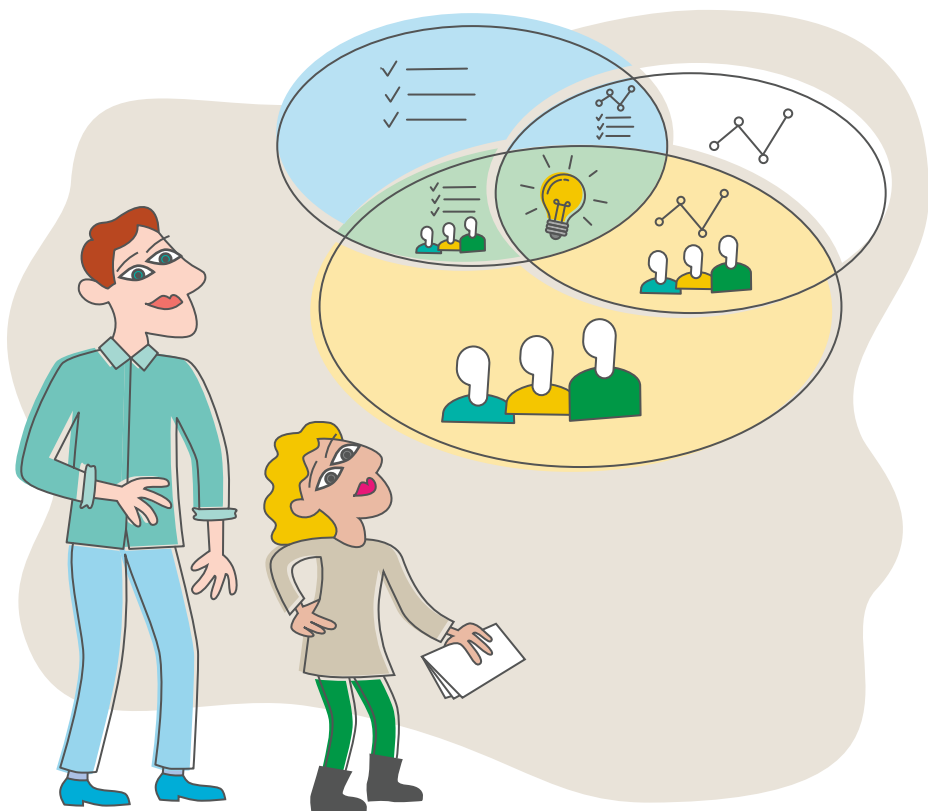
Het uiteindelijke doel is om beter zichtbaar te maken waar leerlingen vastlopen in hun leesbegrip en die inzichten te vertalen naar gerichte ondersteuning. “Zodat leestoetsen zichtbaar maken wat er echt toe doet.”

3 Innovatiethema | **Buiten gewone vakken**

Het volgen van vaardigheden als lezen, burgerschap en digitale geletterdheid stelt scholen voor een opgave. Het meten van vaardigheden die niet in één vak bij één vakdocent beoefend worden, vraagt dat we die vaardigheid over vakken en tijd heen volgen. Hoe verhouden deze meetmomenten zich tot elkaar, en hoe geven ze samen voldoende houvast voor de volgende stap in leren?

Zo'n aanpak roept veel vragen op die traditioneel niet bij losse instrumenten spelen. Vragen die maken dat we grondig moeten kijken naar het systeem van toetsen die een leerling in zijn schoolcarrière maakt. Hoe kunnen we met informatie van veel verschillende opdrachten en verschillende leraren samen komen tot informatie voor en over leren van een leerling?

Deze kernvakken zijn niet uitputtend, maar illustratief voor allerlei vaardigheden waarbij dit speelt en lossen we niet op met een paar losse slimme toetsen. Daarbij komt bij dat deze thema's inhoudelijk vaak volop in beweging zijn op het moment dat onderwijs er al in moet opleiden. Van het meten van diverse vaardigheden in projectonderwijs, het invullen van burgerschapsonderwijs vanuit verschillende levensbeschouwelijke achtergronden, en de razendsnelle ontwikkelingen in digitale vaardigheden; het meten van dit soort vaardigheden doet beroep op onze innovatiekracht in onze expertise van meten.



Onderwijsvraagstukken die in dit thema centraal staan:

- Hoe kunnen toetsen leerlingen en leraren actief betrekken bij het leerproces dat vakoverstijgend plaatsvindt?
- Hoe kunnen we opdrachten en toetsen flexibel ontwerpen zodat ze meebewegen met veranderende onderwijsvisies en actuele maatschappelijke thema's?
- Hoe kun je vanuit datapunten uit allerlei vakken komen tot beoordelingen en beslissingen?
- Hoe kunnen we de validiteit en betrouwbaarheid van toetsen waarborgen wanneer vaardigheden steeds breder en complexer worden én over verschillende vakgebieden heen gemeten moeten worden?

Onze belofte: We bedenken toetsoplossingen die een breed beeld geven en leren en toetsen verbinden met de maatschappij.

Notes to know



In gesprek met Tijmen Poell (innovatieondersteuner) over het verzamelen en gebruiken van leerlinginformatie

Met het project ‘Notes to know’ richten we ons op het ondersteunen van leraren bij het verzamelen en interpreteren van verschillende soorten informatie over leerlingen. Denk aan aantekeningen uit gesprekken of observaties in de klas. Die informatie is er al bij veel leraren, maar wordt nog weinig systematisch benut.

“In het project werken we aan een AI-tool die deze informatie samenbrengt. Een docent kan een map met bestanden uploaden, waarna de AI de inhoud analyseert. Zo ontstaat een rijker en samenhangend beeld van een leerling, met de mogelijkheid om gericht door te vragen over bijvoorbeeld specifieke vaardigheden,” vertelt Tijmen.

Van losse notities naar samenhang

De huidige focus ligt op leraren in het voortgezet onderwijs, met name binnen het havo- en vwo-vak O&O (Onderzoek & Ontwerpen), waar Tijmen zelf lesgeeft. “In de praktijk zien we dat informatie over leerlingen vaak versnipperd wordt vastgelegd. Veel leraren werken nog graag op papier en zijn terughoudend in het structureel vastleggen van datapunten.” Door structuur en overzicht te bieden, wordt het eenvoudiger om gegevens bij te houden en ontstaat er meer ruimte om observaties vast te leggen.

Eerste werkende basis

Een eerste versie van de tool is beschikbaar. Gebruikers kunnen bestanden uploaden en vervolgens in gesprek gaan met de AI over de inhoud. Daarmee ontstaat een eerste vorm van analyse en duiding. Tegelijkertijd is het nog te vroeg om conclusies te trekken. “De komende periode ligt de nadruk op doorontwikkeling. Bijvoorbeeld door het toevoegen van standaard prompts en het ontwikkelen van visualisaties die inzichten beter zichtbaar maken. Daarnaast moet ook de techniek verbeterd worden. Zo moet de snelheid van de reactie van de AI omhoog.”

Samen testen in de praktijk

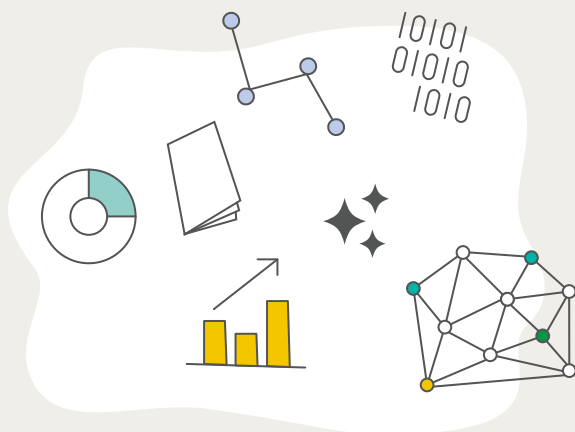
Om het prototype verder te toetsen, worden gesprekken gevoerd met scholen over het gebruik van leerlingdata in de klas. Via Stichting Technasium zijn hiervoor de eerste contacten gelegd. De verwachting is dat deze samenwerking helpt om scholen beter te betrekken bij het testen, terwijl het projectteam zich richt op de inhoudelijke doorontwikkeling. De samenwerking loopt door zolang de externe partner meerwaarde ziet in het prototype.

Aansluiting bij bredere ontwikkelingen

Het project sluit aan bij ontwikkelingen zoals programmatisch toetsen, waarin per competentie informatie wordt verzameld om een rijker beeld van leerlingen te krijgen. Niet één toetsmoment, maar een breder en doorlopend beeld. “Notes to know is in opzet klein en eenvoudig, maar heeft de potentie om breder inzetbaar te worden dan vooraf gedacht. Mogelijk zelfs in verschillende onderwijscontexten.”

Vooruitblik

De volgende stap is het verder verfijnen van de tool, zodat de AI niet alleen informatie bundelt, maar ook actief helpt om patronen en inzichten zichtbaar te maken. “Het doel is om leraren beter te ondersteunen bij het begrijpen van hun leerlingen, op basis van informatie die er al is.”



Meetbaar met portfolio |

Nieuw project

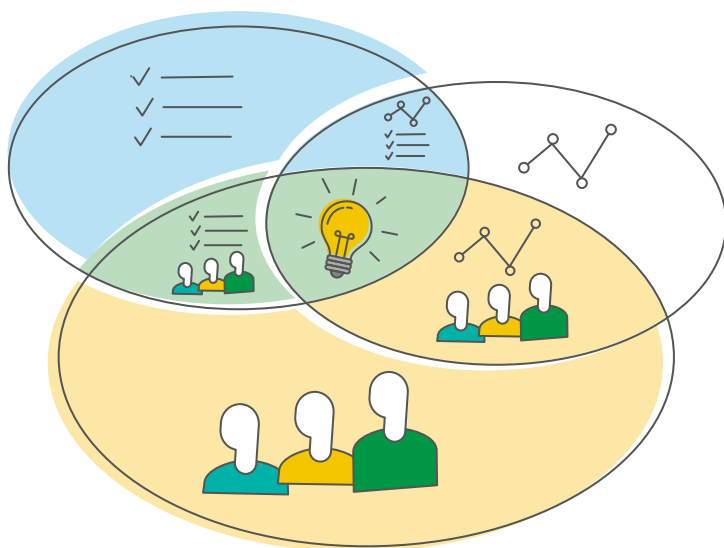
Praktische vaardigheden zijn toegepaste, uitvoeringsgerichte vaardigheden die nodig zijn voor het functioneren in de (toekomstige) beroepscontext en/of zelfredzaamheid in de maatschappij. Ze zijn dus van belang, maar moeilijk meetbaar met standaardtoetsen. Een portfolio geeft leerlingen de kans om werk, reflecties en casussen te bundelen en daarmee hun competenties zichtbaar te maken.

In dit onderzoek ontwikkelen we met evidence-centered design een gebruiksvriendelijk en betrouwbaar raamwerk voor zo'n portfolio. Daarna valideren we ons raamwerk in samenwerking met een school.

Hoe meten we over vakken heen? |

Nieuw project

In het vernieuwde curriculum is meer aandacht voor samenhang tussen leerstof van verschillende vakken. Vaardigheden zoals probleemoplossen,



samenwerken, digitale geletterdheid en burgerschap. Het zijn vaardigheden die je bij uitstek ontwikkelt in complexere opdrachten. Hoe zorg je dat je als leerling, maar ook als team van leraren, dit zorgvuldig ontworpen leerproces over vakken heen volgt en waar nodig met inzichten en reflecties uit andere vakken verrijkt? In dit ontwerp-onderzoek gaan we op zoek naar een instrument dat dienend is aan dit overkoepelende proces.

Notes to know |

Doorontwikkelingsproject

In het projectonderwijs, zoals profielwerkstukken en O&O-opdrachten, geven leraren tussendoor feedback en verzamelen ze allerlei informatie over hun leerlingen: observaties, notities, presentaties, teksten, reflecties en groepsevaluaties. Het overzichtelijk bewaren en benutten van deze losse, ongestructureerde informatie is lastig, terwijl juist de samenhang waardevol inzicht geeft in de prestaties. Vooral wanneer deze informatie ook projectoverstijgend beschikbaar is. In dit project onderzoeken we of AI kan helpen bij het samenvoegen, samenvatten en analyseren van deze bronnen. Zodat leraren en mentoren de verkregen inzichten direct kunnen gebruiken in hun onderwijspraktijk.

Inzicht in burgerschap op school |

Nieuw promotieonderzoek

In dit project onderzoeken we met een breed consortium hoe scholen burgerschapsonderwijs monitoren en welke instrumenten zij daarvoor gebruiken. Dit heeft tot doel om impactvolle manieren van het monitoren van burgerschap te ontwikkelen. We voeren casestudies uit, waarbij we verschillende perspectieven meenemen, zoals die van schoolleiders, leraren en leerlingen. En we analyseren hoe bestaande monitorings-instrumenten worden ingezet, en hoe resultaten worden gebruikt en verwerkt in de kwaliteitszorg. Daarnaast maken we gebruik van evidence-centered design en argumentgericht valideren om monitoringsinstrumenten door te ontwikkelen en te testen. Het onderzoek bestaat uit meerdere fasen, inclusief inventarisatie, ontwerp, beproeving en analyse van kwalitatieve en kwantitatieve data.

4 Innovatiethema | **Balans tussen mens en machine**

De grote hoeveelheid data uit huidige technologische oplossingen biedt veel mogelijkheden om leren zichtbaar te maken, en om leren en toetsen slimmer en efficiënter te maken. Tegelijk is er, misschien meer dan ooit, behoefte aan menselijkheid, duiding en persoonlijk contact. Technologie kan processen personaliseren maar het is de aandacht die iets persoonlijk maakt. En bovendien: niet alle informatie van waarde is te vangen in data. De vraag draait niet om wat er technisch mogelijk is, maar om wat er didactisch en pedagogisch gewenst is.

Om te voorkomen dat de ethische en maatschappelijke discussie over de impact van technologie wordt gekaapt door alleen sceptici of juist door louter enthousiastelingen, vliegen we dit technologische thema juist aan vanuit de waarden van het onderwijs. De uitdaging is vervolgens helder: hoe versterken we die onderwijskundige insteek met technologie, zodat die daadwerkelijk het leerproces ondersteunt. Van slimme instrumenten die leraren relevante informatie opleveren, tot op procesverlichting voor toetsprofessionals, zodat de menselijke kern van beoordelen onverminderd centraal staat.

Onderwijsvraagstukken die in dit thema centraal staan:

- Hoe kunnen we de menselijke aandacht behouden of zelfs vergroten, als we meer digitaal meten en verwerken?
- Hoe kunnen we zorgen dat we meer informatie halen voor didactisch handelen uit de data van onze opgaven?
- Hoe kunnen we technieken gebruiken voor authentiekere leer- en toetservaringen?
- Hoe kunnen we toetsmakers helpen om slimmer hun toetsen te maken, samen te stellen, te verwerken, na te kijken en te analyseren?

Onze belofte: We creëren leeromgevingen waarin digitale innovatie en menselijk contact hand in hand gaan.



Geen checklist, wel het goede gesprek



In gesprek met Max van Haastrecht (onderwijskundig onderzoeker) over verantwoord omgaan met AI in het onderwijs

Waarom dit project?

Volgens Max kent het project twee duidelijke aanleidingen. “Binnen Cito leefde al langer de vraag: hoe kijken wij naar AI en wanneer zetten we het wel of niet in? Tegelijk kregen

we diezelfde vraag steeds vaker van buitenaf, bijvoorbeeld van leraren. Zij wilden weten: hoe denken jullie hierover?” Die vragen vormden de basis voor het project. Niet om iets volledig nieuws te onderzoeken, benadrukt Max, maar om bestaande kennis en ervaringen samen te brengen. “We hebben de afgelopen jaren veel geleerd over AI in relatie tot toetsing en onderwijskundig meten. Dit project draait vooral om het onder woorden brengen van die lessen.”

Van wetenschappelijk artikel naar praktijkgerichte publicatie

Samen met collega's schreef Max een wetenschappelijk artikel waarin die inzichten zijn uitgewerkt. Al snel volgde de vraag of er ook een toegankelijker, minder technische versie kon komen. “Die wens hadden we zelf ook al,” vertelt Max. “In februari is een praktische, Nederlandstalige versie in het vakblad Examens verschenen.” Ook is er behoefte aan een checklist. “Leraren vragen vaak om een checklist: als ik dit doe, zit ik dan goed? Maar het is niet altijd zinnig om als externe partij een checklist ‘op te leggen’, omdat de context altijd mee moet wegen in hoe AI op een waardevolle manier ingezet kan worden,” zegt Max. “Elke school is anders, elke context is anders. Wat wij wél kunnen bieden, is een manier van denken en een set vragen die je jezelf als docent of school kunt stellen.”

Geen one-size-fits-all

Dat uitgangspunt bepaalt ook de koers voor het komende jaar. “Goed advies geven kan alleen als je de context kent: de school, de docent, de klas en het doel. We kunnen niet zeggen dat een bepaalde AI-toepassing overal en altijd werkt. Zelfs bij prototypes waar wij zelf enthousiast over zijn, zoals CheckMate en SchrijfBlik, geldt dat.” De belangrijkste les die het team wil meegeven aan scholen: er bestaat geen

oplossing die altijd werkt, maar ook geen oplossing die nooit werkt. “AI moet niet zomaar overal geïntegreerd worden zonder dat er goed over is nagedacht.”

Van denken naar doen

Omdat scholen vaak weinig tijd hebben om dit soort vragen uitgebreid te verkennen, wil het team ondersteuning bieden in de vorm van een workshop. “We willen het gesprek echt samen aangaan,” legt Max uit. “Niet alles platslaan tot een checklist, maar scholen helpen om hun eigen vragen, context en doelen scherp te krijgen.”

In 2026 wil het team hiermee aan de slag, samen met scholen. “We willen onderzoeken wat werkt, hoe je zo’n gesprek het beste kunt voeren en hoe we dat samen kunnen doorontwikkelen. Het doel is dat er aan het einde van het jaar iets ligt wat scholen daadwerkelijk kunnen blijven gebruiken.”

Wat AI blootlegt: de relatie tussen docent en leerling

Een belangrijke opbrengst van het project zit volgens Max op een minder zichtbaar, maar cruciaal niveau. “In discussies over goede toetsing gaat het vaak over de kwaliteit van de toets of de prestaties van de leerling. De relatie tussen docent en leerling krijgt traditioneel weinig aandacht.” Met de komst van AI verandert dat. “AI-toepassingen raken direct aan vertrouwen, autonomie en privacy. Neem SchrijfBlik: het geeft leraren inzicht in het schrijfproces van leerlingen, wat kan helpen bij beoordeling. Tegelijk kan een leerling het gevoel krijgen: vertrouwt mijn docent mij wel? Mag ik nog mijn eigen ding doen?” Die spanning heeft impact. “Als een docent niet goed weet wat leerlingen met AI doen, kan dat leiden tot wantrouwen. Leraren geven aan: dit opent mijn ogen. Of juist: ik geloof niet dat leerlingen het zó veel gebruiken. Er ontstaat soms een mismatch tussen verwachtingen van leraren en leerlingen.”

Oproep aan scholen

Tot slot doet Max een oproep. “We zijn op zoek naar scholen die met ons willen meedenken: scholen met een concrete vraag of casus, of scholen die willen bijdragen aan het ontwikkelen van zo’n werkvorm.” Een voorbeeld? “Hoe ga je om met het meten van schrijfvaardigheid nu leerlingen AI kunnen gebruiken bij schrijfopdrachten? Dat zijn precies de vragen waar we samen met scholen mee aan de slag willen.”

Test Assembly Studio |

Nieuw project

Het samenstellen van toetsen vraagt om een zorgvuldige balans tussen onderwerpen, moeilijkheidsgraad en versiesamenstellingen, zodat toetsen vergelijkbaar en betrouwbaar zijn. Met Test Assembly Studio ondersteunen we toetsmakers bij dit complexe proces. De tool helpt toetsen samen te stellen terwijl alle randvoorwaarden worden bewaakt. In dit project maken we de tool geschikt voor verschillende toetsketens en ontwikkelen we algoritmen die automatisch toetsvoorstellen genereren die aan alle eisen voldoen. De toetsmaker kan deze voorstellen vervolgens naar eigen inzicht aanpassen. Zo wordt het samenstellen van toetsen efficiënter en minder foutgevoelig.

Vraag & Vind: vraaggestuurde rapportages? |

Nieuw project

Scholen beschikken over steeds meer data en uitgebreide rapportages uit nationale toetsen, maar het is niet altijd eenvoudig om daar de juiste inzichten uit te halen. Waardevolle informatie blijft soms onbenut en leraren verschillen in de vragen die zij met deze data willen beantwoorden. In dit project onderzoeken we of een slimme AI-agent hierbij kan helpen. Deze tool stelt leraren in staat om in 'gewone' taal

vragen te stellen over hun leerlingen, bijvoorbeeld of hun leesvaardigheid is verbeterd, en toont direct de bijbehorende grafieken of tabellen.

Zo staat niet de data zelf, maar de vraag van de leraar centraal, en verkennen we hoe nationale datasets beter kunnen bijdragen aan de dagelijkse onderwijspraktijk.

SchrijfMonitor |

Nieuw project

Uit meerdere onderzoeken blijkt dat de schrijfvaardigheid van leerlingen onder druk staat. Generatieve AI biedt kansen én uitdagingen, voor zowel het aanbieden van als het nakijken van schrijfoopdrachten. Binnen het project SchrijfBlik richten we ons op de impact van AI op het schrijfproces, en in het project SchrijfMonitor onderzoeken we of AI kan ondersteunen bij de beoordeling ervan. Dit kan leiden tot slimmere en schaalbare monitoring van schrijfvaardigheid. Slimme monitoring draagt niet alleen bij aan een beter overzicht van deze belangrijke vaardigheid, maar potentieel levert het ook diagnostische informatie op voor leerkrachten over wat leerlingen al beheersen en waar nog oefening nodig is.

Slimme nakijkoplossingen innovatielijn | *Doorontwikkeling innovatielijn*

Het nakijken van open vragen kost docenten veel tijd en is vaak lastig om betrouwbaar uit te voeren, terwijl deze vragen belangrijk zijn om verschillende vaardigheden valide te meten.

Nieuwe AI-technieken kunnen leerling-antwoorden analyseren en bieden kansen voor toets- en itemanalyse en nakijkondersteuning, maar het is nog onduidelijk welke toepassingen in welke context het meest effectief en verantwoord zijn. Dit project bouwt

voort op eerder onderzoek en vergroot onze kennis over slimme, breed inzetbare en schaalbare nakijkoplossingen.

We ontwikkelen algoritmes en software, en onderbouwen deze met wetenschappelijk onderzoek, zodat docenten sneller en betrouwbaarder kunnen nakijken.

SchrijfBlik | *Doorontwikkelingsproject*

Steeds meer leerlingen gebruiken AI bij schrijfopdrachten, waardoor de resultaten minder betrouwbaar zijn als maat voor schrijfvaardigheid. In 2024 ontwikkelden we SchrijfBlik, een



prototype dat het AI-gebruik van leerlingen tijdens het schrijfproces zichtbaar maakt. In 2025 onderzochten we het op verschillende middelbare scholen, waarbij we ons focusten op het docent- en leerling perspectief. Docenten kregen beter inzicht in hoe leerlingen AI inzetten, terwijl leerlingen verdeelde meningen hadden over dat docenten hun schrijfproces konden inzien en het mee konden nemen in de beoordeling. In 2026 ontwikkelen en onderzoeken we SchrijfBlik verder in samenwerking met de onderwijspraktijk.

Waarde van AI |

Doorontwikkelingsproject

AI speelt een steeds grotere rol in onderwijs en toetsing. Dit roept veel vragen op over wanneer AI waarde toevoegt en wanneer niet. In 2025 hebben we in een publicatie uitgelegd hoe je aan de hand van drie kern-elementen kunt onderzoeken of inzet van AI waarde toevoegt: de onderwijskundige context, de conceptualisering van validiteit en het sociale contract tussen leerling en docent. In 2026 vertalen we dit artikel naar een praatplaat. Hiermee gaan we in gesprek met docenten in het VO en MBO om best practices op te halen en te onderzoeken wat zij nodig hebben om de juiste keuzes te maken. Dit vertalen we naar een workshop waarmee we vanaf

eind 2026 onderwijsinstellingen ondersteunen om onderbouwde keuzes te maken.

Q-Insight | *Doorontwikkelingsproject*

Q-Insight ondersteunt toetsontwikkelaars met data-gedreven suggesties om toetsen eerlijker en inclusiever te maken. In 2025 is de tool al door meerdere instellingen ingezet voor de screening van examens en toetsen op voorspelbaarheden, complexe zinnen en cultuurgebonden taal. In 2026 onderzoeken we welke aanvullende checks nodig zijn om de toetskwaliteit maximaal te versterken. Concreet gaan we op zoek naar kwaliteitscontroles die het gebruik van Q-Insight praktischer en betrouwbaarder maken, zodat de tool vanzelfsprekend wordt in het constructieproces. Daarnaast zorgen we voor veilige verwerking van toetsdata, zodat toetsontwikkelaars de screenings met vertrouwen kunnen uitbesteden en zich kunnen focussen op de inhoudelijke afwegingen die leiden tot de beste toetsen.

CoCo | *Doorontwikkelingsproject*

In dit project onderzoeken we hoe AI het menselijke proces van het creëren van toetsvragen het beste kan ondersteunen. Toetsvragen maken is vaak tijdrovend en vraagt veel expertise.

Wij richten ons daarom op uiteenlopende makers en hun specifieke behoeften. Voortbordurend op het prototype CoCo (Constructie Co-pilot) werken we aan een platform waarmee we het constructieproces ondersteunen. We onderzoeken of en hoe verschillende functies wenselijk zijn voor verschillende gebruikers: voor leraren en voor toetsprofessionals. Centraal staat hoe AI en menselijke deskundigheid elkaar versterken, zodat toetskennis via technologie breder en effectiever inzetbaar wordt in uiteenlopende constructieprocessen.

Beter nakijken én sneller klaar? - De adoptie van AI-nakijkhulp voor open vragen in het VO | *Nieuw promotieonderzoek*

Hybride nakijksystemen waarin de docent samenwerkt met AI-nakijkhulp bij open vragen kunnen de nakijktijd verkorten en de consistentie van beoordeling vergroten. Toch worden deze systemen in het voortgezet onderwijs nog weinig gebruikt. In dit promotieonderzoek onderzoeken we welke factoren de adoptie van AI-nakijkhulp beïnvloeden. We starten met een literatuurreview, waarin we op basis van het innovatieframework van Rogers relevante adoptiefactoren identificeren. Vervolgens voeren we een mixed-methods veldstudie uit op vier scholen

om te verkennen hoe deze in de praktijk werken. Tot slot onderzoeken we in een kwantitatieve veldstudie onder 120 docenten of AI-nakijkhulp daadwerkelijk leidt tot tijdswinst en consistentere beoordeling.

Geavanceerde spraaktechnologie en learning analytics voor leesonderwijs op maat | Eindfase promotieonderzoek

Sinds 2022 werken we binnen het ASTLA-project aan het verbeteren van het meten van de technische en vloeiende leesvaardigheid via nieuwe AI-technieken. Samen met de Radboud Universiteit, Universiteit Twente en Expertisecentrum Nederlands ontwikkelden en valideerden we een meetinstrument dat leesproblemen bij leerlingen tijdig(er) opspoort en leesonderwijs op maat mogelijk maakt. Verder hebben we didactisch relevante leesprofielen uit de data onttrokken, en hebben we onderzocht of het mogelijk is om op basis van deze profielen automatisch feedback te genereren. In 2025 en 2026 integreren we deze ontwikkelingen in een applicatie die de leesproblemen en kwaliteiten van leerlingen automatisch weergeeft, zodat leerkrachten hun leesinstructies op maat kunnen aanbieden.

Schrijven beoordelen met oog voor proces én AI



In gesprek met Marjolein Jongbloed (onderwijskundig onderzoeker) over het prototype SchrijfBlik dat schrijfvaardigheid anders zichtbaar maakt.

Waarom dit project?

Leerlingen gebruiken AI steeds vaker bij het schrijven van teksten. Dat roept een belangrijke vraag op: wat meet je nog als je alleen het eindproduct beoordeelt? Vanuit die ontwikkeling is SchrijfBlik ontstaan. We onderzoeken hoe product- en procesbeoordeling bij schrijfopdrachten gecombineerd kunnen worden, mét ruimte voor het gebruik van AI.

Volgens Marjolein is het uitgangspunt helder: “AI wordt veel gebruikt door leerlingen. Verbieden is niet de oplossing. Je wilt juist begrijpen hoe leerlingen het gebruiken en hoe je daar in je beoordeling recht aan kunt doen.” SchrijfBlik biedt dat inzicht: het wordt voor leraren inzichtelijk hoe leerlingen AI gebruiken bij een schrijfopdracht.

Met het prototype de klas in

SchrijfBlik is uitgebreid getest in verschillende klassen en scholen in het voortgezet onderwijs. Leraren vonden het waardevol om inzicht te krijgen in hoe leerlingen AI gebruiken tijdens het schrijven. Tegelijk schrokken ze soms wel van dat inzicht. “Sommige leraren zagen hoe makkelijk leerlingen tekst van AI overnemen. Dat maakte indruk,” vertelt Marjolein.

Het enthousiasme onder leraren over de inzet van AI bij het schrijven was achteraf minder groot dan vooraf verwacht, ook bij leraren die bij aanvang zeer enthousiast waren. Zij hadden het idee dat SchrijfBlik leerlingen aanzette tot het meer gebruiken van AI, terwijl leerlingen aangaven dat dit niet zo was. De meeste leraren kozen er zelfs voor om de opdrachten niet te beoordelen. Een deel stelde de vraag of schrijfopdrachten weer vaker op papier en in de klas zouden moeten plaatsvinden. Volgens Marjolein speelt mee dat juist taalvakken zijn onderzocht: “Dat zijn precies de vakken waarin AI makkelijker ingezet kan worden om delen te vertalen of delen van de opdracht te laten schrijven.”

Ook zagen we duidelijke verschillen tussen de opdrachten in hoeveel leerlingen kopiëren vanuit AI. Het blijkt dus moeilijk om een opdracht te ontwerpen voor deze vakken waarbij de leerling AI gebruikt op een manier waarop het leerproces wordt ondersteund.

Hoe kijken leerlingen ernaar?

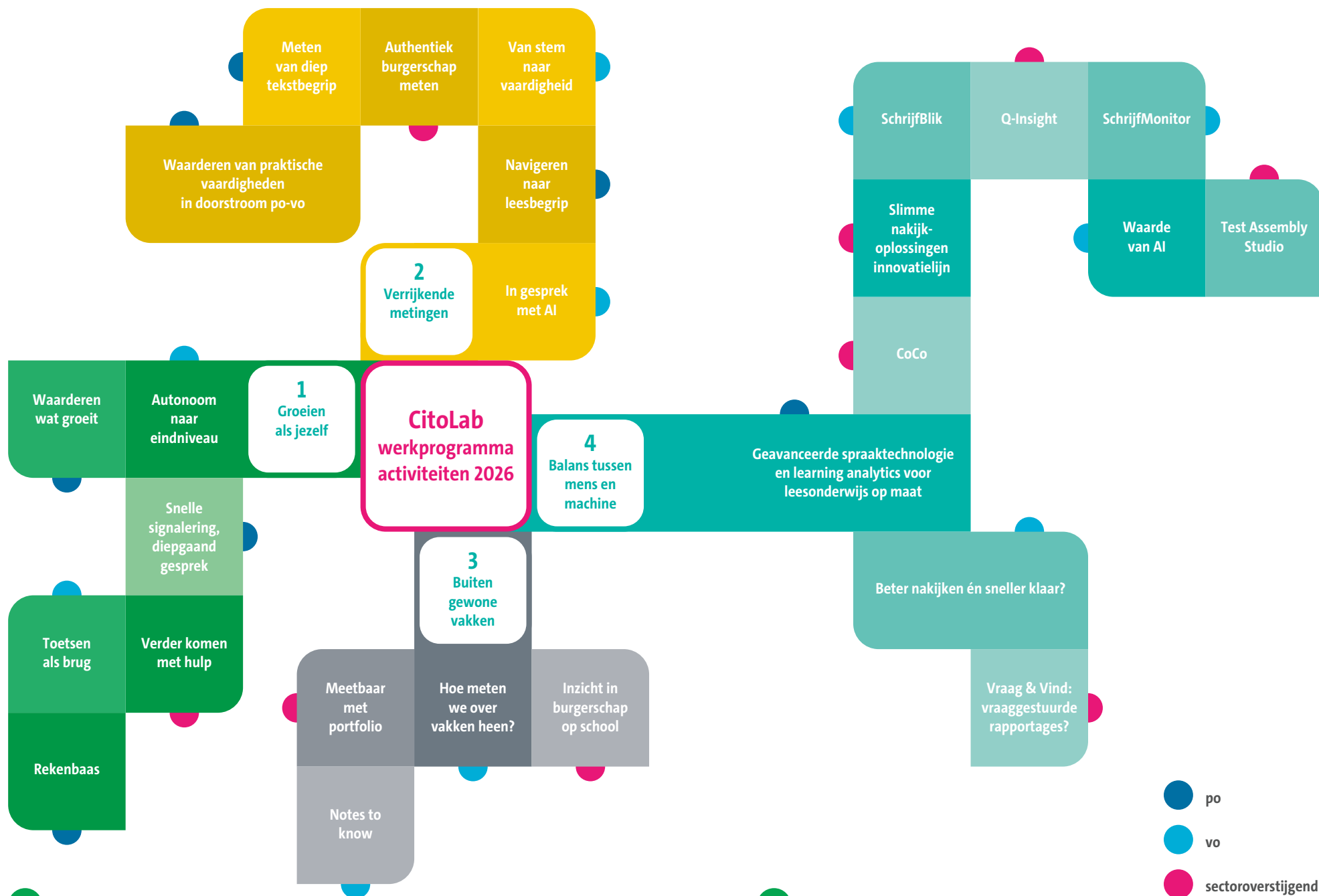
Een deel van de leerlingen vindt het prettig dat leraren beter zicht krijgen op hun schrijfproces en op momenten waarop ze vastlopen. Anderen ervaren dat juist als onprettig en wijzen op privacy. Zij werken liever zoals ze gewend zijn, mét AI, maar zonder dat hun proces volledig inzichtelijk is. Wat Marjolein opviel, is hoe genuanceerd leerlingen over AI spreken. “Sommigen vinden dat AI helemaal niet thuishoort in het onderwijs. Maar de meerderheid ziet het als hulpmiddel, voor inspiratie en ideeën, of zelfs als iets dat gewoon bij hun toekomst hoort.” Leerlingen blijken bovendien opvallend open in hun feedback. “Ze zeggen wat ze denken. Dat levert veel bruikbare inzichten op.”

Vaardigheden en misvattingen rond AI

Leerlingen blijken moeite te hebben met het formuleren van goede prompts. Tegelijk geven leerlingen vaak aan dat ze al veel weten over AI en daar weinig extra over willen leren. “Waarschijnlijk weten ze niet wat ze nog niet weten,” zegt Marjolein. “Daar ligt nog een duidelijke ontwikkelkans.”

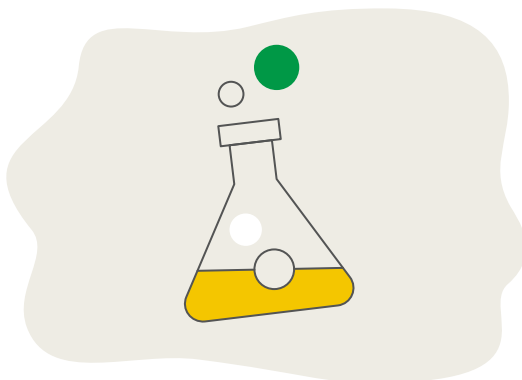
Weerbarstige praktijk, waardevolle betrokkenheid

Het traject bevestigt dat praktijkonderzoek altijd nieuwe ontwerproudes vraagt. Een prototype dat technisch werkt, is nog geen eindproduct voor de klas. “Testen in de praktijk betekent automatisch opnieuw ontwerpen,” zegt Marjolein. Dat is wat we nu doen; we ontwikkelen met de praktijk een verbeterde versie van SchrijfBlik. Daarnaast onderzoeken we hoe schrijfoopdrachten moeten worden ontworpen zodat AI leerlingen ondersteunt in plaats van het leerproces overneemt. Daarmee ontwikkelen we handvaten voor leraren om straks geschikte opdrachten te ontwerpen en te beoordelen.



Tot slot

Dank voor het lezen van het werkprogramma 2026. Ook dit jaar zetten we ons in voor onderwijs waarin iedere leerling een eerlijke kans krijgt om zich te ontwikkelen. Dat doen we door te onderzoeken, te innoveren en samen te werken met diverse partners.



Heb je een idee voor een innovatieproject, zie je kansen voor samenwerking of wil je met ons in gesprek over de toekomst van leren, ontwikkelen en toetsen? Neem dan contact met ons op via citolab@cito.nl.

Samen maken we van 2026 een jaar waarin we kennis omzetten in impact en betekenisvolle stappen zetten voor het onderwijs van morgen.

Colofon

Dit is een uitgave van Stichting Cito, gevestigd te Arnhem.
juli 2026

Tekst

Eline Veerman
Emmelien van der Scheer
CitoLab

Illustraties

Media Support, Cito

Vormgeving en druk

Media Support, Cito
Proces Support, Cito





Cito
Amsterdamseweg 13
6814 CM Arnhem
Postbus 1034
6801 MG Arnhem
T (026) 352 11 11
www.cito.nl

