

LLM-gestuurde trainingsapplicatie voor sociale werkplaatsen

Maxim Peeters
Universiteit Hasselt

Hasselt, Belgium
maxim.peeters@student.uhasselt.be

Abstract—In sociale werkplaatsen zijn traditionele PDF-handleidingen vaak een aanzienlijke barrière voor operatoren met beperkte ervaring of een taalbarrière. De complexe structuur, wisselende lay-out en ingewikkelde teksten maken zelfstandig werken uitdagend en kunnen soms de veiligheid in gevaar brengen. Deze paper beschrijft een nieuw webplatform, ontwikkeld in samenwerking met de sociale werkplaats Open Atelier, dat deze statische instructies omzet in interactieve, digitale modules. Het platform biedt directe ondersteuning via een ingebouwde AI-assistent en bevat een quizsysteem om de opgedane kennis te testen. Door deze combinatie van live AI-hulp en actieve herhaling willen we de veiligheid en de kwaliteit van het werk verhogen, de prestaties verbeteren en de zelfstandigheid van de operatoren versterken.

I. INTRODUCTIE

Maatwerkbedrijven, voorheen beter bekend als sociale werkplaatsen, spelen een cruciale rol op de Vlaamse arbeidsmarkt. Sinds de inwerkingtreding van het Decreet Collectief Maatwerk (2013, van kracht sinds 2019) is hun primaire doelstelling het bieden van zinvolle tewerkstelling en maatschappelijke integratie aan personen met een afstand tot de arbeidsmarkt [1]. Deze medewerkers vormen een heterogene groep, variërend van anderstalige nieuwkomers en laaggeletterden tot personen met een fysieke of cognitieve beperking [2]. Voor hen betekent een aanstelling in de sociale economie niet enkel financiële onafhankelijkheid, maar ook een kans op maatschappelijke participatie en praktische taalverwerving [3].

In de praktijk verloopt de instructie op de werkvloer echter niet altijd drempelvrij. Veel technische werkinstructies en veiligheidsvoorschriften worden nog steeds aangeboden als statische, papieren documenten. Voor medewerkers met een taalbarrière of leesproblemen vormt dit een aanzienlijke drempel. Wanneer instructies niet correct worden begrepen, heeft dit een negatieve invloed op de productiviteit en stijgt het risico op gevaarlijke situaties. De Vlaamse maatwerksector tracht hierop in te spelen door sterker in te zetten op digitale operatorondersteuning [4].

Om dit probleem aan te pakken, worden handleidingen in de praktijk soms handmatig vereenvoudigd. Dit is echter een uiterst tijdrovend proces waarbij vaak cruciale achtergrondinformatie verloren gaat. Hoewel formele opleidingsmomenten ondersteuning bieden bij de opstart, ontbreekt naderhand vaak een systematische controle op het begrip van de instructies. Passieve leermethoden, zoals het louter doorlezen van tek-

sten, blijken in de praktijk onvoldoende effectief. Onderzoek naar het 'testing effect' toont daarentegen aan dat actieve zelftoetsing, bijvoorbeeld via een korte quiz, de kennisretentie aanzienlijk verbetert ten opzichte van passieve studiemethoden [5].

Kennisretentie is bovendien een continu proces. De klassieke vergetenscurve van Ebbinghaus (1885), recent gerepliceerd door Murre en Dros [6], illustreert dat nieuw verworven informatie snel wegzakt zonder herhaling. Bij routinematige werkzaamheden treedt daarnaast het risico op routineblindheid op. Wanneer operatoren handelingen veelvuldig herhalen, kan er onbewust gewenning optreden. In de veiligheidsliteratuur staat dit fenomeen bekend als de *normalisation of deviance*: het proces waarbij kleine afwijkingen van de veiligheidsvoorschriften gaandeweg als norm worden geaccepteerd [7]. Gespreide herhaling (*spaced repetition*) biedt hier een oplossing. Door de leerstof op strategische momenten te herhalen, wordt de vergetenscurve telkens onderbroken en afgevlakt (zie Figuur 1). Hierdoor behouden medewerkers hun focus, wordt een hogere kwaliteitsstandaard nagestreefd en verbetert de veiligheid op de werkvloer op de lange termijn.

Om deze uitdagingen het hoofd te bieden, is in dit project een digitaal leer- en instructieplatform ontwikkeld. In plaats van bestaande, visueel geoptimaliseerde handleidingen te vervangen, voegt het ontwikkelde platform een ondersteunende digitale laag toe. Een geïntegreerde AI-assistent ondersteunt medewerkers tijdens de uitvoering door terminologie te vertalen, stappen te vereenvoudigen of te wijzen op veiligheidsrisico's. Daarnaast genereert het platform automatisch quizvragen via generatieve AI. Dit maakt het mogelijk om direct na afronding van een taak het begrip te verifiëren, en ondersteunt via gespreide herhaling de retentie op lange termijn.

Deze paper is als volgt opgebouwd. Eerst bespreken we in Sectie II de bestaande literatuur over digitale operatorondersteuning en slimme leersystemen. Daarna legt Sectie III het ontwerp en de technische architectuur van het platform uit. In Sectie IV behandelen we praktische uitdagingen rond context-injectie, intellectueel eigendom, privacy, en de cognitieve en ethische impact op de werkvloer. We sluiten af met Sectie V, waarin we de belangrijkste conclusies trekken en ideeën voor vervolgonderzoek aandragen.

The Forgetting Curve

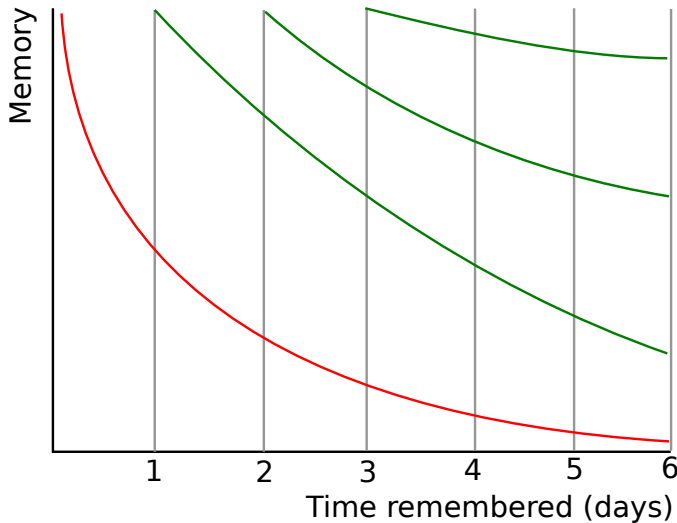


Fig. 1. De invloed van gespreide herhaling (*spaced repetition*) op de vergetenscurve, gebaseerd op het klassieke Ebbinghaus-model (bron: Wikimedia Commons).

II. GERELATEERD WERK

Het idee achter ons platform brengt kennis samen uit digitale hulp voor operatoren, onderwijs technologie en leerpsychologie. Om duidelijk te maken waar ons werk precies over gaat, bespreken we in dit deel eerdere onderzoeken over hulp voor operatoren die de normale werking niet in de war stuurt. Daarnaast kijken we naar hoe we Large Language Models (LLM's) kunnen inzetten om te helpen met taal en om automatisch oefenvragen te bedenken. Tot slot bekijken we hoe we gespreid herhalen op de werkvloer in de praktijk kunnen gebruiken.

A. Aanvullende hulp voor operatoren versus versturende systemen

Omdat er steeds meer verschillende producten worden gemaakt en de werknemers steeds diverser zijn, is goede en passende ondersteuning voor operatoren in de industrie en logistiek erg belangrijk [4]. Veel recent onderzoek kijkt naar heel moderne oplossingen, zoals Augmented Reality (AR), Virtual Reality (VR) of 3D-simulaties. Hoewel die systemen veel voordelen hebben, zijn ze voor kmo's en maatwerkbedrijven vaak te ingewikkeld en te duur [8]. Bedrijven moeten dan namelijk al hun bestaande handleidingen helemaal opnieuw maken. Daarnaast kost de aankoop van hardware, zoals VR-brillen, veel geld. Ook kruipt er te veel tijd in het opleiden van de medewerkers om hiermee te kunnen werken.

In plaats van de bestaande werkwijze volledig te vervangen, stellen we in dit onderzoek een extra laag voor die de huidige manier van werken gewoon behoudt. Bedrijven kunnen hun bestaande papieren handleidingen of PowerPoint-slides, die vaak al jarenlang goed werken, gewoon blijven gebruiken. Ons platform voegt daar alleen een extra digitale stap aan

toe, die helpt wanneer medewerkers moeite hebben met lastige woorden of wanneer ze een andere taal spreken.

Dit past goed bij wat Vanneste en collega's [9], [10] hebben gevonden. Zij laten zien dat digitale werkinstructies ervoor zorgen dat mensen minder fouten maken en beter zelfstandig kunnen werken. Dit geldt vooral voor operatoren met een verstandelijke beperking of mensen die de taal niet goed spreken [9]. Die onderzoeken focussen zich alleen vaak op heel dure oplossingen, zoals projecties op de werktafel of slimme brillen. Vanneste en collega's [10] geven zelf ook toe dat kleinere bedrijven die systemen vaak niet willen, omdat ze te duur zijn om aan te schaffen en veel tijd kosten om te onderhouden. Met ons platform proberen we die problemen op te lossen. Door de hulp via een webapp op een gewone tablet aan te bieden, maken we het veel makkelijker toegankelijk. Zo krijgt de medewerker nog steeds de juiste mentale ondersteuning, zonder de hoge kosten.

B. LLM's als leerassistent en voor het bedenken van vragen

Binnen het onderwijs en bij trainingen worden Large Language Models (LLM's) steeds belangrijker. In dit project focussen we ons vooral op het gebruik van die modellen om teksten makkelijker te maken voor anderstaligen en om automatisch oefenvragen te bedenken [11].

Het zelf bedenken van goede quizvragen kost begeleiders echter heel veel tijd. Zeker in een werkomgeving waar taken snel veranderen, is dit lastig te combineren met hun normale werk. Systemen op basis van LLM's kunnen hierbij helpen door zelfstandig vragen te bedenken over de leerstof, inclusief het juiste antwoord en een paar foute antwoordopties. Uit een recente studie van Mitroulias en Sioutas [11] blijkt dat moderne LLM's hier erg goed in zijn. Wel blijft de betrouwbaarheid een probleem: de modellen verzinnen soms foute dingen (ze 'hallucineren'), waardoor ze verkeerde vragen kunnen maken. In een werkomgeving waar veiligheid heel belangrijk is, is dat een groot risico.

Om dat risico te verkleinen, wordt er in onderzoeken vaak gekozen voor een menselijke controle. Hierbij werkt het AI-model eigenlijk als een assistent die alvast een eerste versie van de vragen maakt. Daarna kijkt een begeleider of expert de vragen na, past ze waar nodig aan en keurt ze goed voordat de werknemer ze te zien krijgt. Op die manier combineer je de snelheid van het automatisch genereren met de zekerheid van een menselijke check.

C. Kennis beter onthouden met gespreid herhalen op de werkvloer

Iets maar één keer uitleggen is meestal niet genoeg om ervoor te zorgen dat mensen het lang onthouden. Om te voorkomen dat nieuwe informatie snel weer vergeten wordt (wat Ebbinghaus ook beschrijft met zijn vergetenscurve [6]), werkt gespreid herhalen heel goed. Onderzoek laat ook zien dat je dingen veel beter onthoudt door jezelf te testen, bijvoorbeeld met een korte quiz. Dat werkt veel beter dan de informatie gewoon opnieuw lezen [5], [12].

Veel bestaande leer-apps gebruiken hiervoor vaste, automatische schema's [12]. In een maatwerkbedrijf of in de industrie zijn die strenge schema's vaak onhandig. Mensen werken daar in ploegendiensten, wisselen van taken en de planning verandert vaak. Onderzoek laat zien dat een opbouwend schema het beste werkt [12], [13]. Hierbij krijg je eerst snel een herhaling (bijvoorbeeld na een week) om de leerstof op te frissen, en daarna zit er steeds meer tijd tussen de quizen (bijvoorbeeld een maand). Op de werkvloer is het daarom handig om die schema's wat losser in te plannen. Zo kun je rekening houden met hoe druk het is of hoe gevaarlijk een bepaalde taak is. Op die manier zorg je ervoor dat werknemers niet blind worden voor risico's, terwijl het toch goed in te plannen blijft [7].

III. IMPLEMENTATIE

Om de hierboven beschreven theorie over extra ondersteuning en betere retentie te vertalen naar een werkende applicatie, hebben we een volledige webapp ontwikkeld. We wilden vooral dat de originele werkinstructies behouden bleven, dat het AI-model naadloos geïntegreerd werd, en dat we op een ethische manier kunnen testen wat de mensen onthouden hebben. De volledige programmacode van dit project staat online op GitHub¹.

A. Werking en technologiekeuzes

Het platform is opgedeeld in een moderne front-end en een robuuste back-end die via een REST API met elkaar communiceren (zie Figuur 2). Het systeem is ontworpen om efficiënt om te gaan met zware processen, zoals het ontleden van PDF-bestanden en de continue communicatie met externe AI-modellen.

Er is bewust gekozen voor een back-end in Java met het Spring Boot-framework, in plaats van alternatieven zoals Node.js of Python. In de context van een maatwerkbedrijf is een uiterst stabiele en veilige data-affhandeling een harde vereiste. Spring Boot biedt een robuuste enterprise-architectuur voor complexe integraties en veilige opslag van leerresultaten in de PostgreSQL-database [14].

Aan de kant van de gebruiker (de front-end) is gekozen voor een webapplicatie gebouwd met React en TypeScript. React zorgt ervoor dat de interface (zoals de AI-chatbot en de quiz) direct reageert zonder de pagina opnieuw te laden. Dit is cruciaal om een vloeiende, interactieve ervaring te garanderen, zelfs op de vaak oudere en tragere tablets die op de werkvloer circuleren. De visuele opmaak is bewust minimalistisch gehouden met Tailwind CSS, zodat operatoren niet afgeleid raken.

Tot slot is het volledige systeem in containers geplaatst met Docker. Door gebruik te maken van Docker Compose worden de drie hoofdcomponenten (de React front-end, de Spring Boot back-end en de PostgreSQL database) als één samenhangend geheel opgestart. Dit garandeert een geïsoleerde en reproduceerbare architectuur. Voor een maatwerkbedrijf zonder uitgebreide IT-afdeling betekent dit een enorme tijds winst: het

platform is nagenoeg 'plug and play' op te zetten op eender welke server, zonder dat men zich zorgen hoeft te maken over conflicterende softwareversies of afhankelijkheden van het besturingssysteem.

B. User-Centric Design en Toegankelijkheid

Bij de implementatie van de front-end stond de eindgebruiker centraal. Aangezien de doelgroep binnen maatwerkbedrijven sterk varieert op gebied van cognitieve en digitale vaardigheden, was een focus op *user-centric design* (mensgericht ontwerpen) cruciaal. De applicatie maakt gebruik van een hoog contrast, responsieve navigatieknoppen en heldere iconografie. Hierdoor ervaren operatoren met bijvoorbeeld een motorische of visuele kwetsbaarheid zo weinig mogelijk fysieke barrières bij het navigeren door hun instructies.

Daarnaast is het design zo opgebouwd dat cognitieve overbelasting wordt vermeden. In plaats van alle mogelijke functionaliteiten en menu's tegelijk te tonen, is de interface bewust minimalistisch gehouden. De interactieve AI-chatbot en de periodieke herhalingsquizen worden pas visueel aangeboden op de momenten dat ze didactisch relevant zijn voor de gebruiker. Deze stapsgewijze presentatie van leerelementen sluit naadloos aan bij de principes van *scaffolding* binnen het onderwijs. Het hoofddoel van de implementatie was immers niet om louter complexe functionaliteiten toe te voegen, maar om een intuïtief hulpmiddel te bouwen dat de operationele zelfstandigheid van de werknemer versterkt zonder af te leiden.

C. Upload-pipeline en Documentverwerking

De eerste stap in het proces is het aanmaken van leermodules. Een beheerder logt in op het persoonlijke dashboard (zie Figuur 3) en laadt de originele PDF-werkinstructie op in het systeem.

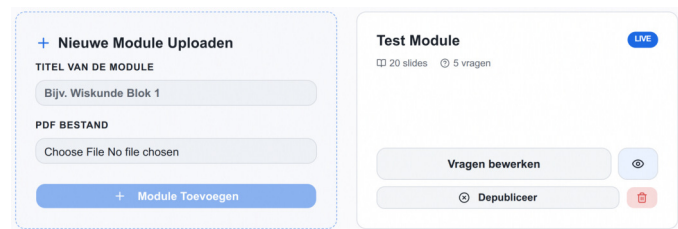


Fig. 3. Het overzichtelijke beheerdersdashboard waar alle leermodules beheerd en geupload kunnen worden.

Omdat het een expliciete vereiste was om de originele instructies ongewijzigd te tonen, is er een automatische data-pipeline gebouwd (zie Figuur 4). Wanneer de PDF geüpload wordt, stuurt de front-end een aanvraag naar de back-end. De server maakt direct een nieuw Module-record aan in de database en start met de verwerking van het document via Apache PDFBox.

Tijdens deze verwerking gebeuren er twee gelijktijdige acties per pagina (slide). Enerzijds wordt de ruwe tekstlaag

¹<https://github.com/Maxim-Peeters/werkplek-leren>

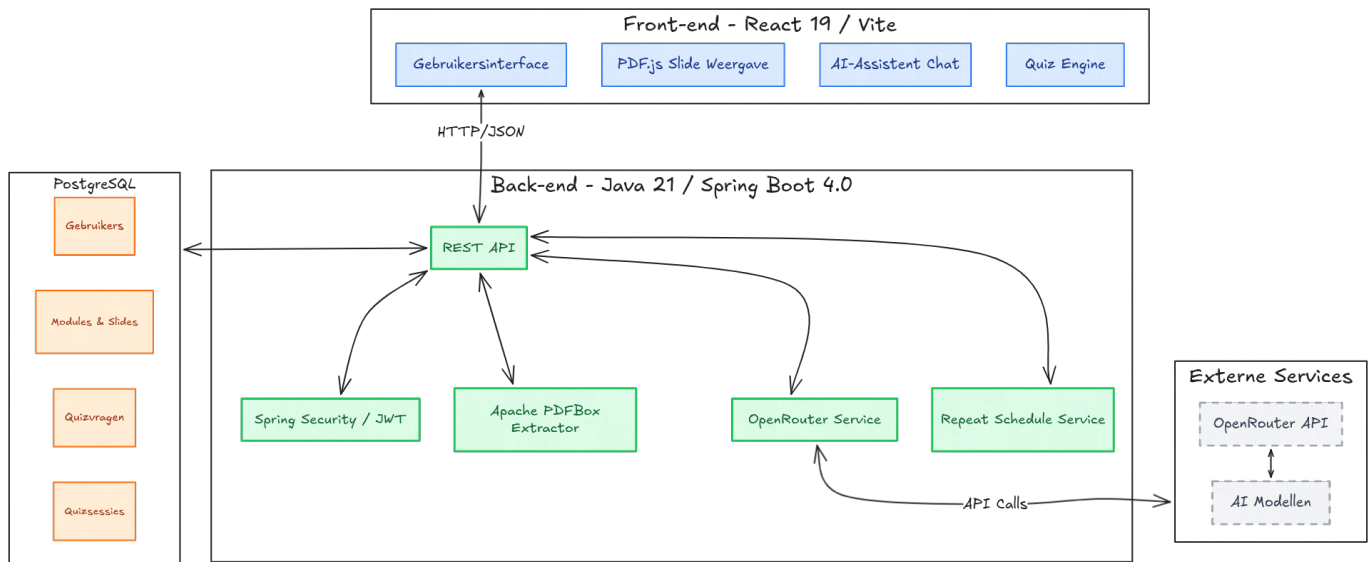


Fig. 2. High-level architectuur van het systeem, met de scheiding tussen front-end, back-end en externe AI-services.

geëxtraheerd; deze tekst vormt de basiscontext voor de latere AI. Anderzijds wordt de visuele weergave van de pagina gerenderd tot een JPG-afbeelding, welke lokaal op de back-end bewaard wordt. Bij dit renderproces via Apache PDF-Box is specifiek gelet op de balans tussen visuele kwaliteit en bestandsgrootte. Door een efficiënte render-resolutie in te stellen, blijven de originele technische tekeningen scherp genoeg voor verwerking door artificiële intelligentie en voor de operator, terwijl de opslagruimte beperkt blijft en de laadtijden op de tablets geminimaliseerd worden. Als laatste stap wordt de database gestructureerd: voor elke slide wordt een record bewaard met het volgnummer, de geëxtraheerde tekst en het afbeeldingspad. Zo ontstaat er een feilloze link tussen het originele beeld en de tekstuele data.

Hoewel het lokaal opslaan van de gerenderde afbeeldingen op de back-end voor een proof-of-concept efficiënt is, is er bewust gekozen om de originele instructies statisch te houden in plaats van interactieve media te ondersteunen. Aangezien de primaire focus van de applicatie ligt op het verbeteren van de taalbeheersing en het testen of de operatoren de theoretische stof daadwerkelijk beheersen, volstaat een heldere visuele weergave. Toch vormt deze lokale opslag een belangrijk architecturaal aandachtspunt: bij verdere opschaling kan het hosten van grote volumes mediabestanden op de applicatieserver leiden tot merkbare prestatieproblemen (*performance issues*). In een productieomgeving is het daarom sterk aanbevolen om deze afbeeldingen in een gespecialiseerde *storage bucket* (zoals Amazon S3) onder te brengen om de server te ontlasten [15].

D. Hybride Quizgeneratie via het Dashboard

Zodra de slides in de database zijn opgeslagen, kan de beheerder via het dashboard (zie Figuur 3) de bijbehorende vragenlijst beheren. Dit proces verloopt bewust niet volledig

automatisch, om zo de eindcontrole bij de expert te houden. De beheerder heeft de flexibiliteit om handmatig specifieke vragen toe te voegen, of om de ingebouwde AI-generatie in te schakelen. Door het gewenste aantal vragen in te geven en de generatie te initiëren, maakt de back-end via de OpenRouter API [16] contact met externe AI-modellen om direct een gepaste set vragen te genereren (voor een uitgebreidere discussie rondom de actuele modelkeuze, zie Sectie III-F).

Voor deze specifieke taak is het essentieel om een AI-model te selecteren dat niet alleen tekst begrijpt, maar idealiter ook de slides visueel kan interpreteren (multimodale *vision support*). Een instructieslide kan immers voornamelijk uit foto's of technische tekeningen bestaan, met minimale tekst. Het geselecteerde model analyseert de slides en retourneert een set van logische meerkeuzevragen, inclusief afleidingsopties (*distractors*).

Om de veiligheid op de werkvloer te waarborgen en AI-fouten te voorkomen, hanteert het systeem een *expert-in-the-loop* flow. De nieuw gegenereerde quiz staat initieel altijd op DRAFT. Pas nadat de beheerder de gegenereerde vragen in het dashboard handmatig heeft nagekeken en goedgekeurd (zie Figuur 5), mag de module de status LIVE krijgen.

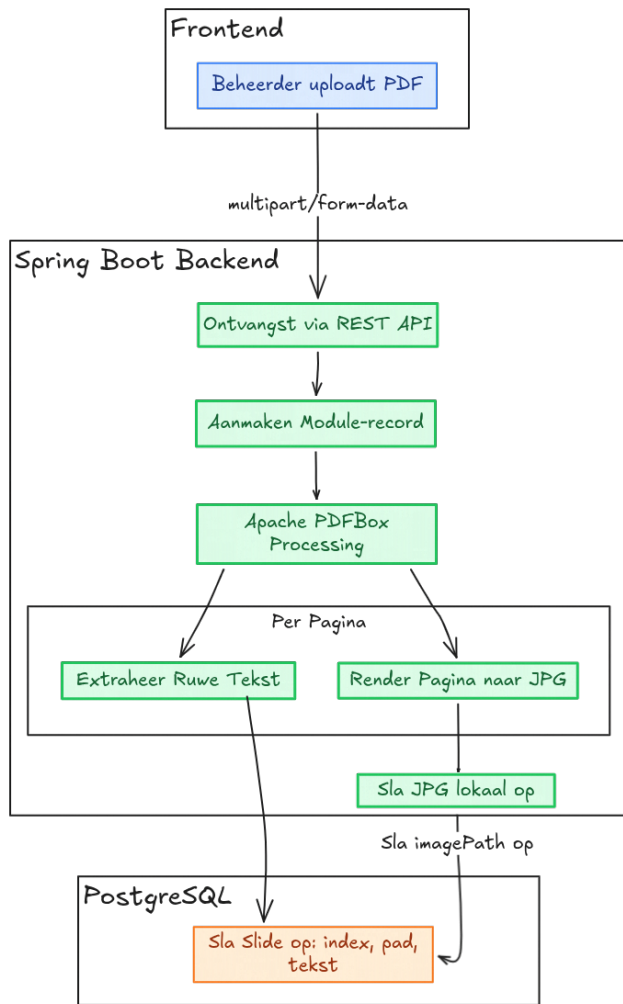


Fig. 4. Schematische weergave van de PDF upload-flow, van front-end tot de PostgreSQL database.

Terug naar Beheerderspaneel

Quiz Vragen: Test Module

1 GENEREER AI OPSLAAN

ACTIEVE STUDENTEN GEDETECTEERD

Vraag 1

Verwijderen

VRAAGTEKST

Je maakt een verlichting voor LINKS. Welke mal gebruik je?

ANTWOORDOPTIES

- ANTWOORD 1: De mal voor RECHTS
- ANTWOORD 2: De mal voor LINKS (JUST)
- ANTWOORD 3: Je hebt geen mal nodig
- ANTWOORD 4: Een eigen mal

Fig. 5. De 'expert-in-the-loop' interface: beheerders valideren en corrigeren AI-gegenereerde vragen alvorens ze in productie gaan.

E. Interactieve PDF-viewer en AI-Chatbot

Nadat de module LIVE staat, krijgt de operator via de tablet toegang. Het platform haalt de eerder gegenereerde JPG-afbeeldingen op en toont de originele werkinstructies via een interactieve slideviewer. De interface van deze weergave is ontworpen met toegankelijkheid in het achterhoofd. Grote, duidelijke knoppen en een minimalistisch design zorgen ervoor dat medewerkers met een visuele of motorische beperking probleemloos door de handleiding kunnen navigeren. Naast de mogelijkheid om per slide eigen notities toe te voegen, bevindt zich steeds de AI-chatbot om vragen over de huidige werkinstructie direct te beantwoorden (zie Figuur 6).

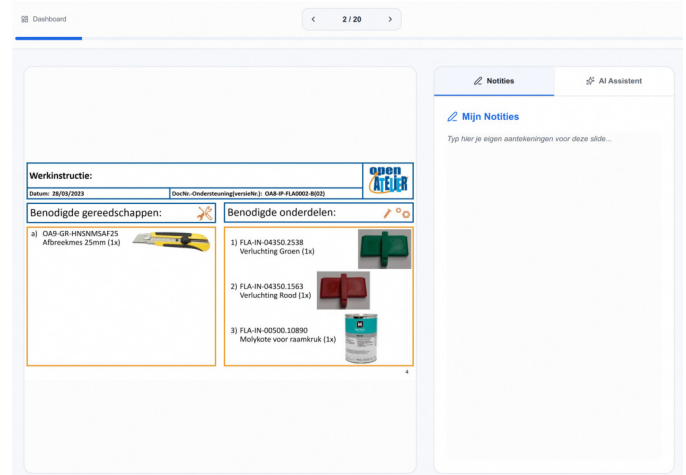


Fig. 6. De weergave voor de operator: een originele instructieslide met daarnaast de contextbewuste AI-chatbot.

De kracht van deze chatbot ligt in de veelzijdige context-injectie. Enerzijds voegt de back-end op de achtergrond telkens alle eerder geëxtraheerde tekst van de actieve slide toe aan de verborgen prompt (de opbouw van deze *system prompts* en *pedagogical steering* wordt in detail behandeld in Sectie III-G). Deze tekst dient als essentiële verduidelijking wanneer de gebruiker een vraag stelt. Anderzijds krijgt de AI-assistent ook de gerenderde afbeelding van de slide mee, waardoor het model daadwerkelijk "ziet" wat erop staat. Deze techniek kan beschouwd worden als een gerichte vorm van Retrieval-Augmented Generation (RAG). In plaats van een abstracte zoekopdracht in een vector database uit te voeren, injecteert het systeem hier direct de actieve slidecontext. Uit recent onderzoek blijkt dat een dergelijke multimodale context-injectie de precisie van AI-assistenten in industriële omgevingen aanzienlijk verhoogt [17]. Hierdoor kan artificiële intelligentie feilloos visuele details interpreteren, zoals de exacte vorm of kleur van een onderdeel. Figuur 7 toont hier een voorbeeld van: de operator vraagt naar de kleur van het mes op de instructieslide, en doordat artificiële intelligentie vision-capaciteiten heeft en de slide daadwerkelijk ziet, kan het de vraag correct beantwoorden met "Geel". De assistent 'weet' exact naar welke stap de operator kijkt en kan daardoor

uiterst gerichte uitleg verstrekken, zonder dat de werknemer de situatie zelf hoeft te beschrijven.



Fig. 7. Een detailweergave van de AI-chatbot, die dankzij multimodale vision-capaciteiten visuele vragen over de instructieslide kan beantwoorden.

F. Modelkeuze voor Quiz en Chatbot

Dankzij de OpenRouter integratie kan het platform dynamisch schakelen tussen verschillende AI-modellen, afhankelijk van de taak. Dit sluit aan bij recente academische inzichten rondom 'model routing', waarbij men dynamisch de optimale balans zoekt tussen kwaliteit, latentie en resourcekosten [18]. Elke API-aanroep kent immers een specifieke *cost per token*, wat de architectuur dwingt tot een efficiënte modelkeuze:

Voor de **quizgeneratie** is een geavanceerd model met brede context- en *image support* vereist. Omdat dit een asynchroon proces is dat enkel door beheerders wordt gestart, heeft kwaliteit voorrang op snelheid. Momenteel wordt het platform hiervoor getest met Google's Gemma 3 27B IT. De voornaamste reden voor deze modelkeuze is dat de Gemma-modellen uiterst kostenefficiënt zijn, wat ze perfect maakt voor de huidige testfasen, terwijl ze toch een acceptabele balans bieden tussen semantische nauwkeurigheid en redeneerprestaties.

Voor de **AI-chatbot** daarentegen is een minimale responstijd (lage latency) vereist, aangezien vertragingen de vlotte werking voor operatoren kunnen verstoren. Omdat er bovendien tientallen vragen per uur kunnen passeren, moet de operationele kost zo laag mogelijk blijven. Om die reden maakt de chatbot momenteel gebruik van het lichtere Google Gemma 3 12B IT-model.

Dankzij de modulaire architectuur (via OpenRouter) is het systeem echter volledig *future-proof*: naarmate de functionele vereisten evolueren of wanneer het budget het toelaat, kan er naadloos worden overgeschakeld naar krachtigere, intelligentere modellen (hoewel dit uiteraard gepaard zal gaan met een hogere *cost per token*).

G. Geavanceerde Prompt Engineering en Context-Injectie

Om ervoor te zorgen dat de AI-modellen zich bij elke API-call veilig, strikt en accuraat gedragen, is er binnen de back-end aanzienlijke aandacht besteed aan *prompt engineering*. Zowel voor de quizgeneratie als de chatbot wordt een dynamische samenstelling van instructies en context opgebouwd (zie Figuur 8).

Hoewel beide AI-toepassingen steunen op de geëxtraheerde tekst en afbeeldingen van de PDF (via multimodale context-injectie), verschillen hun *system prompts* wezenlijk qua doel en techniek. Bij de asynchrone generatie van quizvragen ligt de focus hoofdzakelijk op machinale leesbaarheid en feitelijke accuraatheid. Hier dwingen harde constraints het model om uitsluitend een strikt JSON-formaat te retourneren (structured output formatting) volgens het opgelegde JSON-schema. Negatieve instructies verbieden het model nadrukkelijk om onzinnige meta-vragen te verzinnen over bijvoorbeeld paginanummers of titels.

De prompt voor de live chatbot draait daarentegen niet om formele parsing, maar om veilige en begrijpelijke mens-machine-interactie. Hier krijgt het model een actieve *persona* opgelegd: een "geduldige en veilige begeleider". Lange zinnen worden afgestraft en taal op A2-niveau wordt afgedwongen. Het meest cruciale verschil is echter de toevoeging van *pedagogical steering* [19]. Wanneer een operator via de chat om hulp vraagt tijdens een quiz, instrueert de back-end het model via de verborgen prompt om **nooit** direct het antwoord te verklappen. In plaats daarvan wordt het model gedwongen tot *scaffolding* (de Socratische methode): het geeft gerichte hints of stelt een wedervraag om het actieve leereffect te behouden.

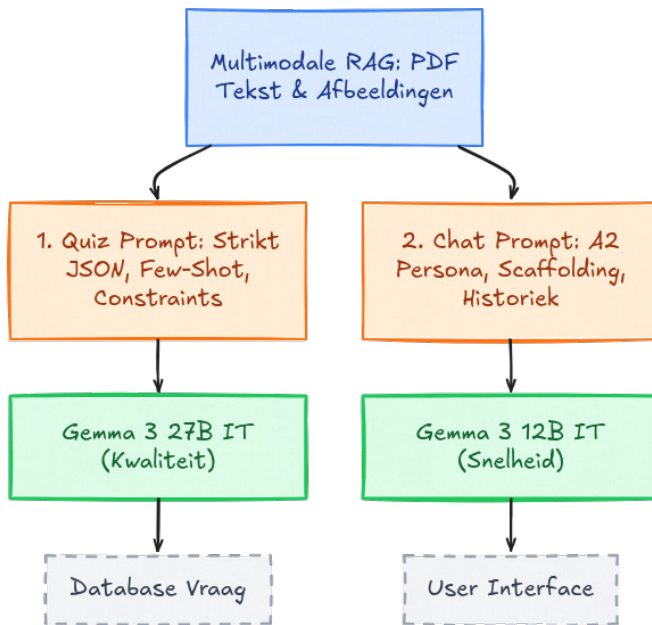


Fig. 8. Architectuur van de dynamische prompts. Beide stromen gebruiken de multimodale context, maar verschillen fundamenteel in sturing en modelkeuze.

H. Retentieschema's Instellen

Het eindstation van de applicatieflow vormt de RepeatScheduleService. Terwijl een operator aan de slag gaat, krijgt deze op periodieke momenten een actieve herhalingsquiz gepresenteerd (zie Figuur 9). Zodra de operator deze quiz volledig heeft doorlopen, bepaalt de RepeatScheduleService exact wanneer de opgedane kennis opnieuw getoetst moet worden volgens het principe van gespreide herhaling (*spaced repetition*), aangezien actieve *retrieval practice* en herhaling cruciaal zijn voor kennisbehoud op de lange termijn [20].

Fig. 9. Een operator die actief een meerkeuzequiz oplost; de score wordt enkel gebruikt voor het retentie-algoritme en niet voor beoordeling.

De timing van deze herhalingen is door de beheerder per individuele module in te stellen (zie Figuur 10). In de praktijk wordt de eerste retentie-sessie bewust op korte termijn ingepland (bijvoorbeeld de volgende dag of na één week) om te garanderen dat cruciale basiskennis meteen na het eerste contact wordt opgefrist. De daaropvolgende retentiemomenten hebben telkens een langer interval (bijvoorbeeld pas na een maand, en vervolgens om de drie maanden).

Fig. 10. Instellen van retentieschema's (spaced repetition) per leermodule via het dashboard.

Belangrijk om te benadrukken is dat het platform enkel de herhalingsmomenten activeert; begeleiders krijgen bewust **geen** overzichten te zien van specifieke fouten of scores van werknemers. De test scores voeden enkel het herhalingsalgoritme, waarmee de psychologische veiligheid op de werkplaats gegarandeerd blijft.

IV. DISCUSSIE EN REFLECTIE

Tijdens de ontwikkeling van dit interactieve leerplatform kwamen er niet alleen technische, maar ook praktische en ethische vraagstukken naar voren. Het is immers belangrijk dat een applicatie niet alleen goed werkt, maar ook echt veilig en bruikbaar is op de werkvloer. In deze sectie bespreek ik de belangrijkste uitdagingen en ontwerpkeuzes.

A. Behoud van Origineel Materiaal en IP

Een van de eerste problemen was dat bedrijven niet graag al hun bestaande pdf-handleidingen opnieuw willen aanpassen o.w.v. de tijdsinvestering. Ook was het belangrijk om het intellectueel eigendom (IP) en de verantwoordelijkheid over de originele veiligheidsvoorschriften te behouden. Als een AI-model een instructie herschrijft en deze blijkt fout te zijn, ontstaat er een juridisch en operationeel risico. Daarom is er gekozen voor een transparante 'overlay'-aanpak. Dit betekent dat de originele pdf altijd exact hetzelfde blijft, inclusief veiligheidsstempels, en het AI-model hier slechts als een onzichtbare laag overheen ligt. Dit bleek een erg pragmatische oplossing te zijn.

B. De Rol van de Begeleider (Expert-in-the-loop)

Hoewel AI veel werk uit handen kan nemen, ben ik via een *expert-in-the-loop* systeem tegemoetgekomen aan de onvoorspelbaarheid van AI-modellen. Doordat een begeleider de automatisch gegenereerde vragen eerst moet goedkeuren, vermijden we dat hallucinaties van het model tot onveilige situaties leiden. Het systeem is zo gebouwd dat het AI-model fungeert als een slimme assistent voor de begeleider, en nooit als eindbeslissers over de leerstof. In een omgeving waar fysieke veiligheid vooropstaat, is deze menselijke controlelaag momenteel absoluut onmisbaar.

C. Kwaliteit van AI en Bronmateriaal

Tijdens het testen werd duidelijk dat de accuraatheid van artificiële intelligentie sterk afhankelijk is van de kwaliteit van de ingeladen pdf (het bekende *Garbage In, Garbage Out* principe) [17]. Als een bedrijf oude of onduidelijk gescande documenten aanlevert, heeft artificiële intelligentie veel moeite om de tekst correct te begrijpen. Daarnaast merkte ik dat de huidige multimodale modellen nog steeds beperkingen hebben. Ze kunnen weliswaar een foto van een tang of schroef herkennen, maar hebben nog erg veel moeite met het correct "lezen" van complexe elektronische schema's of 3D-CAD tekeningen [10]. Voor erg technische assemblages stoot de technologie voorlopig nog op haar limieten.

D. Gebruikerservaring en Adoptie

De doelgroep in een sociale werkplaats heeft vaak een kwetsbaar profiel, waardoor psychologische veiligheid en toegankelijkheid erg belangrijk zijn. Als artificiële intelligentie te lange of ingewikkelde zinnen genereert, kan dit zorgen voor cognitieve overbelasting [8], [9]. Daarom heb ik de interface bewust zo eenvoudig mogelijk gehouden. Een ander belangrijk punt was het voorkomen van stress bij de operatoren. Als werknemers het gevoel krijgen dat hun quizscores worden gebruikt om hen af te rekenen, zullen ze het systeem vermijden. Daarom zijn de scores strikt gescheiden van prestatie-evaluaties door de begeleiders; ze dienen uitsluitend om het algoritme voor gespreide herhaling op een positieve manier te sturen.

E. Economische Haalbaarheid

Een laatste praktisch punt is de werking via cloud-API's zoals OpenRouter. Hoewel dit erg handig is voor de ontwikkeling, brengt het in de praktijk ook nadelen met zich mee. Je betaalt namelijk per opgevraagde prompt (*cost per token*), wat bij honderden actieve gebruikers op een fabrieksvloer kan oplopen. Bovendien eist deze architectuur een stabiele internetverbinding. Een logische volgende stap is dan ook het onderzoeken van lokale AI-modellen die op een eigen server kunnen draaien, om zo de kosten te verlagen en onafhankelijk te worden van externe internetdiensten [18].

V. CONCLUSIE EN TOEKOMSTIG WERK

In dit individueel project heb ik een LLM-gestuurd leerplatform ontwikkeld om de cognitieve en taalkundige ondersteuning in sociale werkplaatsen te verbeteren. Door te kiezen voor de eerder besproken 'overlay'-aanpak, kunnen bestaande werkinstructies heel eenvoudig worden uitgebreid met meer-talige (via vertaling naar de moedertaal) en contextbewuste AI-hulp, zonder de originele bestanden te veranderen.

Ik heb in dit systeem ook actief geprobeerd om inzichten uit de leerpsychologie te verwerken. Door gebruik te maken van actieve herinnering en gespreide herhaling (*spaced repetition*), wil dit platform voorkomen dat operatoren door routine blind worden voor de veiligheidsvoorschriften. De privacy en de eigenwaarde van de werknemer staan hierbij centraal, aangezien de tool vooral dient als een persoonlijk vangnet en niet als een evaluatiemiddel voor het management.

Hoewel dit platform al een mooie basis vormt, zijn er nog verschillende pistes voor de toekomst:

A. Evolutie van Modelselectie

Voor de huidige testfase is er gewerkt met goedkopere AI-modellen zoals Google Gemma 3, wat prima volstond. Het platform is echter flexibel gebouwd. Zodra het nodig is (bijvoorbeeld voor erg complexe of visueel uitdagende taken), kan er dankzij OpenRouter met een simpele aanpassing overgeschakeld worden naar geavanceerdere modellen zoals GPT-4o of Claude 3.5 Sonnet.

B. Dynamische Quizfunctionaliteit en Personalisatie

De quizzen zouden in de toekomst nog interactiever gemaakt kunnen worden. Om te vermijden dat een werknemer na verloop van tijd gewoon de antwoorden van buiten leert, zou het systeem bij een nieuwe poging telkens andere vragen kunnen genereren. Daarnaast zou een "bladwijzer"-functie handig zijn, waarmee operatoren moeilijke stappen of vragen zelf kunnen opslaan om later opnieuw te bestuderen. Verder kan de personalisatie per operator nog aanzienlijk worden verbeterd door een slim algoritme te integreren. Dit algoritme zou zich automatisch kunnen afstemmen op de individuele vooruitgang, het aantal gemaakte fouten, en de specifieke tijd die een operator per instructie nodig heeft. Op basis hiervan kan de moeilijkheidsgraad of de frequentie van vragen dynamisch aangepast worden, en kunnen handmatige bladwijzers gebruikt worden om gericht herhalingsoefeningen te genereren.

C. Multimodale Validatie van Open Vragen

Op dit moment werkt het platform enkel met meerkeuzevragen. Met de kracht van huidige AI-modellen zou het interessant zijn om ook open vragen toe te laten. Het AI-model zou dan kunnen evalueren of een ingetypt of ingesproken antwoord voldoende dicht bij de waarheid ligt, en kan gerichte feedback geven op wat er mist in de uitleg.

D. Kwantitatieve Veldstudie

Dit onderzoek richtte zich vooral op de ontwikkeling en de proof-of-concept. Om echt te weten of het systeem werkt zoals gehoopt, is er een langdurige test in de praktijk nodig. Alleen door een uitgebreide veldstudie kunnen we meten of operatoren daadwerkelijk veiliger of zelfstandiger gaan werken over een periode van enkele maanden.

E. Persoonlijke Reflectie

Terugkijkend op dit project, was het bouwen van dit platform een grote en boeiende uitdaging. Op technisch vlak heb ik me moeten verdiepen in technologieën die nieuw voor me waren, zoals de Spring Boot back-end, de React front-end en het gebruik van Docker-containers. Het was een leerzaam proces om deze tools goed met elkaar en met de AI-modellen te laten communiceren. Daarnaast gaf de probleemstelling, die aansluit bij de praktijk van Open Atelier, me een veel beter beeld van de maatschappelijke kant van informatica. Het vertalen van complexe code naar een intuïtieve interface voor een kwetsbare doelgroep dwong me om echt vanuit de gebruiker te denken. Dit project heeft me dus niet alleen betere programmeerskills opgeleverd, maar heeft me ook geleerd hoe belangrijk het is om verantwoorde en gebruiksvriendelijke software te bouwen.

DANKWOORD

Graag wil ik mijn promotoren, Prof. dr. Eva Geurts en Prof. dr. Kris Luyten, bedanken voor hun waardevolle sturing en feedback gedurende dit traject. Daarnaast wil ik Open Atelier bedanken voor het aanreiken van de interessante probleemstelling en mogelijkheid om met hun werkinstructies aan de slag te gaan.

REFERENCES

- [1] Vlaamse Overheid, "Decreet betreffende maatwerk bij collectieve inschakeling," Vlaamse Codex, 2013, decreet van 12 juli 2013, in werking getreden 1 januari 2019. Geconsolideerde versie geldig op 26 mei 2026. Geraadpleegd op 21 mei 2026. [Online]. Available: <https://codex.vlaanderen.be/Portals/Codex/documenten/1023214.html>
- [2] Departement Werk & Sociale Economie, "Beleidsnota sociale economie 2024-2029: Inclusief en duurzaam ondernemerschap in vlaanderen," Vlaamse Overheid, Brussel, België, Tech. Rep., 2024, geraadpleegd op 22 mei 2026. [Online]. Available: <https://www.vlaanderen.be/publicaties/beleidsnota-2024-2029-sociale-economie>
- [3] H. Lenaerts, B. Audenaert, J. Baki, P. Cools, S. Marchal, M.-L. Mulayi, W. Neelen, L. Van Herreweghe, W. Van Lancker, and B. Cantillon, "Sociale economie en armoede. De impact van de sociale economie op armoede in vlaanderen (social): evaluatie en vooruitblik," Centrum voor Sociaal Beleid Herman Deleeck, Universiteit Antwerpen, Antwerpen, België, Onderzoeksrapport, 2024, geraadpleegd op 23 mei 2026. [Online]. Available: <https://publicaties.vlaanderen.be/view-file/65306>
- [4] Agentschap Innoveren & Ondernemen, "Begeleidingstrajecten ondersteunen shift naar industrie 4.0," VLAIO Nieuws, Apr. 2022, gepubliceerd op 20 april 2022. Geraadpleegd op 24 mei 2026. [Online]. Available: <https://www.vlaio.be/nl/nieuws/begeleidingstrajecten-ondersteunen-shift-naar-industrie-40>
- [5] H. L. Roediger, III and J. D. Karpicke, "Test-enhanced learning: Taking memory tests improves long-term retention," *Psychological Science*, vol. 17, no. 3, pp. 249–255, 2006, geraadpleegd op 25 mei 2026. [Online]. Available: https://colinallen.dnsalias.org/Readings/2006_Roediger_Karpicke_PsychSci.pdf

- [6] J. M. J. Murre and J. Dros, "Replication and analysis of ebbinghaus' forgetting curve," *PLoS ONE*, vol. 10, no. 7, p. e0120644, 2015, geraadpleegd op 26 mei 2026. [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0120644>
- [7] N. Sedlar, A. Irwin, D. Martin, and R. Roberts, "A qualitative systematic review on the application of the normalisation of deviance phenomenon within high-risk industries," *PsyArXiv Preprint*, 2022, geraadpleegd op 27 mei 2026. [Online]. Available: https://osf.io/preprints/psyarxiv/tqphk_v1
- [8] F. Krause, T. Bosch, E. Wilschut, and G. van Rhijn, "Cognitive operator support in the manufacturing industry: Three tools to help smes select, test and evaluate operator support technology," in *Human-Technology Interaction: Shaping the Future of Industrial User Interfaces*. Springer, 2023, pp. 49–82. [Online]. Available: https://doi.org/10.1007/978-3-030-99235-4_3
- [9] P. Vanneste, Y. Huang, J. Y. Park, F. Cornillie, B. Decloedt, and W. Van den Noortgate, "Cognitive support for assembly operations by means of augmented reality: An exploratory study," *International Journal of Human-Computer Studies*, vol. 143, p. 102480, 2020. [Online]. Available: <https://doi.org/10.1016/j.ijhcs.2020.102480>
- [10] P. Vanneste, K. Dekeyser, L. A. Pinos Ullauri, D. Debeer, F. Cornillie, F. Depaepe, A. Raes, W. Van Den Noortgate, and S. Said-Metwaly, "Towards tailored cognitive support in augmented reality assembly work instructions," *Journal of Computer Assisted Learning*, vol. 40, no. 2, pp. 797–811, 2024. [Online]. Available: <https://doi.org/10.1111/jcal.12916>
- [11] D. Mitroulias and S. Sioutas, "A systematic review and bibliometric analysis of automated multiple-choice question generation," *Big Data and Cognitive Computing*, vol. 10, no. 1, p. 35, 2026. [Online]. Available: <https://doi.org/10.3390/bdcc10010035>
- [12] S. H. K. Kang, "Spaced repetition promotes efficient and effective learning: Policy implications for instruction," *Policy Insights from the Behavioral and Brain Sciences*, vol. 3, no. 1, pp. 12–19, 2016. [Online]. Available: <https://doi.org/10.1177/2372732215624708>
- [13] N. J. Cepeda, H. Pashler, E. Vul, J. T. Wixted, and D. Rohrer, "Distributed practice in verbal recall tasks: A review and quantitative synthesis," *Psychological bulletin*, vol. 132, no. 3, p. 354, 2006. [Online]. Available: <https://psycnet.apa.org/doiLanding?doi=10.1037/2F0033-2909.132.3.354>
- [14] O. Novac, D. Ghiurau, M. Novac, C. Gordan, M. Oproescu, G. Bujdosó, and V. O. Mihalca, "Comparison of node.js and spring boot in web development," in *2023 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*. IEEE, 2023, pp. 1–6.
- [15] Amazon Web Services, "Amazon simple storage service (s3) documentation," <https://docs.aws.amazon.com/s3/>, 2023, geraadpleegd op 6 juni 2026.
- [16] OpenRouter, "Openrouter: A unified interface for llms," <https://openrouter.ai/>, 2024, geraadpleegd op 6 juni 2026.
- [17] M. Riedler and S. Langer, "Beyond text: Optimizing rag with multimodal inputs for industrial applications," *arXiv preprint*, 2024. [Online]. Available: <https://arxiv.org/abs/2410.21943>
- [18] C. Varangot-Reille, C. Bouvard, A. Gourru, M. Ciancone, M. Schaeffer, and F. Jacquenet, "Doing more with less: Implementing routing strategies in large language model-based systems: An extended survey," *arXiv preprint arXiv:2502.00409*, 2025. [Online]. Available: <https://arxiv.org/abs/2502.00409>
- [19] S. Sonkar, N. Liu, D. Basu Mallick, and R. G. Baraniuk, "Class: Conversational learning with analytical step-by-step strategies," *arXiv preprint*, 2023. [Online]. Available: <https://arxiv.org/abs/2305.13272>
- [20] J. D. Karpicke and H. L. Roediger, "The crucial role of retrieval practice in long-term retention," *Science*, vol. 319, no. 5865, pp. 966–968, 2008. [Online]. Available: <https://www.science.org/doi/10.1126/science.1152408>