



Professionele Bachelor Toegepaste Informatica



Optimalisatie-Interactie: Educatieve chatbots voor leerlingen met leerstoornissen

Josephine Bouwens

Promotoren:

Thomas Verhoef externe promotor
Sam van Rijn interne promotor

Eureka Expert
Hogeschool PXL Hasselt



Bachelorpaper Academiejaar 2025-2026



Professionele Bachelor Toegepaste Informatica



Optimalisatie-Interactie: Educatieve chatbots voor leerlingen met leerstoornissen

Josephine Bouwens

Promotoren:

Thomas Verhoef externe promotor
Sam van Rijn interne promotor

Eureka Expert
Hogeschool PXL Hasselt



Bachelorpaper Academiejaar 2025-2026

Dankwoord

Graag wil ik enkele mensen bedanken die een belangrijke bijdrage hebben geleverd aan het ontstaan van deze paper.

In de eerste plaats wil ik mijn promotoren, Thomas Verhoef en Sam van Rijn, oprecht bedanken voor hun waardevolle begeleiding, feedback en ondersteuning tijdens dit traject. Hun expertise en betrokkenheid vormden een grote meerwaarde voor dit onderzoek.

Daarnaast wil ik Anny Cooreman bedanken voor het vertrouwen en voor de kans om dit project te realiseren. Ook wil ik Filip Hamakers bedanken voor de interne ondersteuning en zijn betrokkenheid tijdens dit traject.

Verder wil ik de leerlingen van Eureka Onderwijs bedanken voor hun bereidheid om deel te nemen aan de tests. Zonder hun medewerking was dit onderzoek niet tot stand gekomen.

Ook wil ik mijn ouders bedanken voor hun voortdurende steun, vertrouwen en aanmoediging tijdens dit traject. Samen met de begeleiding, feedback en betrokkenheid van alle hierboven genoemde personen heeft dit mij geholpen om deze paper succesvol af te ronden.

Abstract

In deze bachelorproef wordt onderzocht hoe AI-gestuurde educatieve chatbots kunnen worden afgestemd op de noden van leerlingen met leerstoornissen, met een focus op dyslexie en dysfasie. De stage vindt plaats binnen Eureka Expert BV, onderdeel van Eureka Leuven, een toonaangevend Belgisch expertisecentrum die personen met leerstoornissen begeleidt.

Binnen deze context wordt een prototype ontwikkeld van een AI-gedreven chatbot die optreedt als digitale studiebegeleider. Het systeem moet studeren ondersteunen door de leerling te ondervragen en gericht feedback te geven op basis van het studiemateriaal van de gebruiker. Hierbij wordt specifiek rekening gehouden met de noden van leerlingen met dyslexie en/of dysfasie.

Het systeem wordt ontwikkeld in React (TypeScript/Vite). De AI-functionaliteit wordt gerealiseerd via een Node.js/Express-backend met Google Gemini voor tekst- en spraakverwerking. Door gebruik te maken van Retrieval-Augmented Generation (RAG) via Gemini File Search worden antwoorden inhoudelijk afgestemd op het studiemateriaal.

Daarnaast wordt een literatuurstudie uitgevoerd naar informatieverwerking en effectieve leerstrategieën bij leerlingen met dyslexie en/of dysfasie. Op basis hiervan wordt onderzocht hoe een AI-chatbot deze strategieën kan ondersteunen. Daarnaast wordt onderzocht hoe system prompts geoptimaliseerd kunnen worden en wordt een promptstrategie uitgewerkt.

Tot slot wordt een evaluatie uitgevoerd met leerlingen gericht op hun gebruikservaring. Aan de hand van gerichte vragen wordt gepeild naar gebruiksvriendelijkheid, duidelijkheid, motivatie en ondersteuning. De resultaten worden verwerkt in een evaluatiematrix, die toelaat om verbeterpunten te formuleren voor zowel de chatbotinterface als de onderliggende promptstructuur.

Uit de ontwikkeling en evaluatie blijkt dat de AI-gestuurde chatbot potentieel heeft als aanvullende studieondersteuning voor leerlingen met dyslexie en/of dysfasie. De combinatie van Retrieval-Augmented Generation, aangepaste promptstrategieën, eenvoudige taal, herhaling en multimodale interactie zorgt ervoor dat de chatbot beter kan aansluiten bij de noden van deze doelgroep. Doordat de antwoorden gebaseerd worden op het studiemateriaal van de leerling, kan de ondersteuning inhoudelijk relevanter en meer contextgericht worden aangeboden.

De gebruikerstesten tonen aan dat de chatbot door leerlingen als motiverend en behulpzaam kan worden ervaren. Vooral de mogelijkheid om zelfstandig te oefenen, vragen te stellen, uitleg te krijgen en via spraak te communiceren, maakt de toepassing toegankelijker voor leerlingen die moeilijkheden ervaren met lezen, schrijven of taalverwerking. De chatbot biedt bovendien ruimte voor herhaling en directe feedback, wat kan bijdragen aan meer zelfvertrouwen en zelfstandigheid tijdens het studeren.

Tegelijkertijd moet de chatbot worden gezien als een ondersteunend hulpmiddel en niet als vervanging van een leerkracht of begeleider. De resultaten tonen aan dat AI waardevol kan zijn binnen gepersonaliseerde leerondersteuning, op voorwaarde dat de technologie zorgvuldig wordt afgestemd op de doelgroep en ingebed wordt in een veilige, pedagogisch verantwoorde leercontext. Deze bachelorproef toont daarmee aan dat een AI-gestuurde studiebegeleider een veelbelovende aanvulling kan vormen op bestaande digitale leerplatformen voor leerlingen met leerstoornissen.

Inhoudsopgave

Dankwoord	ii
Abstract	iii
Inhoudsopgave	iv
Lijst van gebruikte figuren	vi
Lijst van gebruikte tabellen	vii
Lijst van gebruikte afkortingen	viii
Inleiding	1
I. Stageverslag	2
1 Bedrijfsvoorstelling	2
2 Stageopdracht	3
2.1 Probleemstelling	3
2.2 Hoofddoel	3
2.3 Subdoelstellingen	3
2.4 Methodologieën en technologieën	4
2.4.1 Methodologische aanpak	4
2.4.2 Technologische omgeving	4
2.5 Uitwerking	6
2.5.1 Analysefase	6
2.5.2 Ontwikkeling en implementatiefase	6
II. Onderzoekstopic	18
1 Onderzoeksopdracht	18
1.1 Opdracht	18
1.2 Onderzoeksmethodiek	19
1.3 Doelgroepanalyse	19
1.3.1 Gezamenlijke ontwerpaanpak voor dyslexie en dysfasie	19
1.3.2 Definiëring dyslexie en dysfasie	20
1.3.3 Kenmerken dyslexie en dysfasie	20
1.4 Vertaling naar technologische ondersteuning	22
1.4.1 Lezen	22
1.4.2 Schrijven	23
1.4.3 Taalbegrip en instructieverwerking	23
1.4.4 Geheugen en informatieverwerking	24
1.4.5 Aandacht en concentratie	25
1.4.6 Motivatie en zelfvertrouwen	25

1.4.7 Formulering van gedachten.....	26
1.5 STT en TTS-benchmarking	26
1.5.1 STT.....	26
1.5.2 TTS.....	28
1.6 Systeempromptoptimalisatie	30
1.6.1 Definiëring van systeempromptoptimalisatie.....	31
1.6.2 Prompt engineeringtechnieken	31
1.6.3 Ontwerpenprompt.....	35
1.7 Usertesten	39
1.7.1 Ontwerp van de usertesten	39
1.7.2 Resultaten van de usertesten	40
1.8 De evaluatiematrix	50
1.9 Geschat tokengebruik	51
1.10 Aanbevelingen voor de toekomst	55
1.10.1 Uitwerking van adaptieve systeemprompt	55
1.10.2 De GDPR en AI gebruik bij kinderen	55
1.10.3 juridische kanttekening bij het uploaden van leermateriaal	56
1.10.4 Voorlichting over AI-gebruik bij kinderen.....	56
1.10.5 Lokaal versus via API-gebaseerde implementatie	56
Conclusie	57
Reflectie.....	58
Bronnenlijst	59
Bijlagen	63

Lijst van gebruikte figuren

Figuur 1 Organisatiestructuur Eureka.....	2
Figuur 2 Sequentiediagram van de RAG-flow	8
Figuur 3 Sequentiediagram van de live spraakinteractie	10
Figuur 4 Artificial Analysis Word Error Rate (AA-WER) Index.....	27
Figuur 5 Speed Factor	27
Figuur 6 Price of Transcription.....	27
Figuur 7 Text to Speech Arena Quality ELO	29
Figuur 8 Price	29
Figuur 9 Speed	29
Figuur 10 Het verschil tussen standaard prompting en chain-of-thoughtprompting.....	33
Figuur 11 Weergave rol en doel systeem prompt	37
Figuur 12 Interactie instructies systeemprompt voor chatbot vraag	38
Figuur 13 Interactie instructies systeem prompt voor chatbot antwoord.....	38
Figuur 14 Systeemprompt gedragsturing	39
Figuur 15 Taartdiagram leeftijden deelnemers	40
Figuur 16 Taartdiagram leerjaar deelnemers	41
Figuur 17 Staafdiagram van het aantal deelnemers per leer- of ontwikkelingsstoornissen	41
Figuur 18 Staafdiagram van de meest helpende studiemethoden volgens de deelnemers	42
Figuur 19 Taartdiagram over de duidelijkheid van de chatbot.....	42
Figuur 20 Taartdiagram antwoorden deelnemers vinden studeren met chatbot leuker.....	43
Figuur 21 Taartdiagram van de invloed van de chatbot op de studiemotivatie.....	43
Figuur 22 Staafdiagram van de ervaren verbetering in het tekstbegrip.....	44
Figuur 23 Staafdiagram van de begrijpelijkheid van de chatbotantwoorden	45
Figuur 24 Taartdiagram met de bevindingen over de lengte van de uitleg.....	46
Figuur 25 Taartdiagram met de bevindingen van respondenten over de latency bij textgeneratie	47
Figuur 26 Taartdiagram met de bevindingen van respondenten over de latency bij TTS.....	47
Figuur 27 Taartdiagram van de ervaringen van leerlingen met de spraakherkenning van de microfoonfunctie	48
Figuur 28 Taartdiagram van de ervaringen van leerlingen met de spraakherkenning van de conversatiemodus	48
Figuur 29 Staafdiagram weergave belang van een menselijk klinkende stem bij de chatbot.....	49
Figuur 30 Staafdiagram weergave belang van een Vlaamse stem bij de chatbot	49

Lijst van gebruikte tabellen

Tabel 1 Kenmerken dyslexie	21
Tabel 2 Kenmerken dysfasie	22
Tabel 3 De resultaten van de gebruikerstesten weergegeven in een evaluatiematrix	51
Tabel 4 De tokenkost per gebruikt model	51
Tabel 5 Maandelijkse scenariokostenberekening voor Gemini in tokens	52
Tabel 6 Maandelijkse scenariokostenberekening voor Gemini in dollars	53
Tabel 7 Maandelijkse scenariokostenberekening voor Gemini en Elevenlabs in dollars	53
Tabel 8 De tarieven voor Gemini Native Audio	54

Lijst van gebruikte afkortingen

AA-WER	Artificial Analysis Word Error Rate
ADD	Attention Deficit Disorder
ADHD	Attention Deficit Hyperactivity Disorder
AI	Artificiële intelligentie
API	Application Programming Interface
AVG	Algemene Verordening Gegevensbescherming
BV	Besloten Vennootschap
B2B	Business-to-Business
CORS	Cross-Origin Resource Sharing
CSS	Cascading Style Sheets
DCD	Developmental Coordination Disorder
DLD	Developmental Language Disorder
ELO	Elo-rating
GDPR	General Data Protection Regulation
HTTP	Hypertext Transfer Protocol
IP	Internet Protocol
JWT	JSON Web Token
LLM	Large Language Model
NLP	Natural Language Processing
OWASP	Open Worldwide Application Security Project
RAG	Retrieval-Augmented Generation
REST	Representational State Transfer
SAAS	Software as a Service
SQL	Structured Query Language
STT	Speech-to-Text
TOS	Taalontwikkelingsstoornis
TTS	Text-to-Speech
UI	User Interface
USD	United States Dollar
WER	Word Error Rate
XSS	Cross-Site Scripting

Inleiding

De afgelopen jaren zijn AI-gestuurde chatbots steeds prominenter aanwezig in digitale leeromgevingen. Ontwikkelaars ontwerpen de interactiestijl van die systemen meestal met de gemiddelde leerling voor ogen. Er wordt dan ook uitgegaan van uniforme cognitieve vaardigheden, leerstrategieën en taalniveaus.

Deze veronderstellingen sluiten echter vaak onvoldoende aan bij leerlingen met een leerstoornis. Nochtans bieden deze AI-gestuurde systemen juist voor deze doelgroep aanzienlijke mogelijkheden. AI-systemen ondersteunen leerlingen gepersonaliseerd in tempo, herhaling, multimodale uitleg en onmiddellijke feedback. Hierdoor verlopen leerprocessen effectiever en toegankelijker.

Deze bachelorproef presenteert een op maat gemaakte chatbot ontwikkeld in samenwerking met Eureka Expert BV. Binnen Eureka Expert ligt de focus op professionele vormingen, educatieve publicaties en digitale leeroplossingen waaronder het online leerplatform CleverKids waarop deze bachelorproef voortbouwt.

In het kader van de stageopdracht wordt een chatbot ontwikkeld als verlengstuk van dit leerplatform met het oog op een mogelijke latere integratie. De chatbot fungeert als een digitale studiebegeleider en biedt leerlingen ondersteuning tijdens het leerproces. Hierbij wordt er bijzondere aandacht besteed aan personalisatie voor leerlingen met leerstoornissen, in het bijzonder dyslexie en dysfasie, waarbij de interactie- en communicatiestijl van het systeem wordt afgestemd op hun specifieke noden.

Naast de praktische ontwikkeling omvat deze bachelorproef een literatuurstudie. Hierin wordt onderzoek gedaan naar de specifieke noden van jonge leerlingen met dyslexie en/of dysfasie. Concreet onderzoekt men hoe men de *system prompt* optimaliseert om zo gesprek- en antwoordstrategieën te ontwerpen die aansluiten bij hun leer- en communicatiestijl.

In de laatste fase van het onderzoek wordt een evaluatiemoment georganiseerd met leerlingen. Tijdens deze evaluatie worden zij bevraagd over hun ervaringen met de applicatie. Hierbij wordt specifiek aandacht besteed aan verschillende aspecten van de *user experience*, zoals de gebruiksvriendelijkheid van de interface, de duidelijkheid van de antwoorden en de mate waarin de chatbot hen effectief ondersteunt tijdens het leerproces.

De verzamelde feedback wordt vervolgens systematisch verwerkt in een evaluatiematrix. In deze matrix worden de verschillende evaluatiecriteria naast de reacties en beoordelingen van de leerlingen geplaatst, waardoor de resultaten op een overzichtelijke en gestructureerde manier kunnen worden geanalyseerd. Dit maakt het mogelijk om sterke punten en mogelijke verbeterpunten van het systeem duidelijk te identificeren.

Tot slot worden mogelijke uitbreidingen en vervolgonderzoeken besproken met het oog op verdere optimalisatie van de chatbot.

I. Stageverslag

1 Bedrijfsvoorstelling

De stage wordt volmaakt binnen Eureka Expert BV, een entiteit binnen de overkoepelende organisatie Eureka Leuven. Figuur 1 visualiseert de organisatiestructuur.



Figuur 1 Organisatiestructuur Eureka

Eureka ondersteunt normaal tot hoogbegaafde kinderen, jongeren en volwassenen met een leerstoornis. De organisatie ontstond begin jaren negentig en heeft zich sindsdien ontwikkeld tot een toonaangevend kennis- en expertisecentrum in België.

Binnen Eureka Leuven streven de verschillende organisaties naar één en dezelfde missie en visie: personen met leerstoornissen maximale kansen bieden om hun talenten te ontwikkelen, hun leerproblemen te compenseren en volwaardig te participeren in het maatschappelijk leven. [1]

Eureka Expert focust specifiek op professionele vormingen en studiedagen, educatieve publicaties en digitale leeroplossingen waaronder het online leerplatform CleverKids. Via deze initiatieven draagt de organisatie bij aan de verspreiding van wetenschappelijk onderbouwde kennis en praktische hulpmiddelen voor onderwijsprofessionals, leerlingen, ouders en begeleiders. [2]

Voor de technische ontwikkeling van CleverKids werkt Eureka Expert samen met Panenco, een gespecialiseerd softwareontwikkelingsbureau. Panenco ontwikkelt CleverKids onder leiding van Koen Verschooten en Egwin Avau en ondersteunt Eureka Expert bij de technische realisatie van het platform. Het bedrijf werkt vaker aan innovatieve B2B SaaS-systemen en enterprise AI-oplossingen. [3]

Omdat deze stage zowel expertise rond leerstoornissen als inzicht in applicatieontwikkeling vereist, wordt de stage begeleid en uitgevoerd vanuit Eureka Expert. Daarnaast wordt Thomas Verhoef als externe promotor aangesteld om mee in te staan voor de technische begeleiding van de stage.

In het kader van de stageopdracht wordt een chatbot ontwikkeld die aansluit bij het bestaande leerplatform CleverKids. Aangezien CleverKids is ontwikkeld met React, wordt ook de chatbot binnen hetzelfde technologisch kader gebouwd. Dit bevordert de technische compatibiliteit en maakt een eventuele toekomstige integratie in het platform mogelijk.

Tijdens de stage wordt de chatbot echter als een afzonderlijke toepassing ontwikkeld, zodat de bestaande architectuur van CleverKids niet rechtstreeks wordt beïnvloed en de implementatie gecontroleerd en modulair kan verlopen.

2 Stageopdracht

2.1 Probleemstelling

Binnen het hedendaagse onderwijslandschap verwachten scholen steeds vaker dat leerlingen de leerinhouden zelfstandig verwerken. Hoewel scholen digitale leermiddelen steeds vaker in de klas integreren, blijft gepersonaliseerde begeleiding buiten de lesmomenten vaak beperkt beschikbaar. Leerkrachten beschikken immers over beperkte tijd om individuele ondersteuning te bieden, zeker in heterogene klasgroepen waar leerlingen sterk uiteenlopende noden hebben.

Deze problematiek manifesteert zich nog sterker bij leerlingen met specifieke leerstoornissen, zoals dyslexie of dysfasie. Hoewel er al digitale hulpmiddelen bestaan, bieden deze vaak slechts generieke ondersteuning. Ze zijn doorgaans niet afgestemd op specifieke handboeken of leerplannen en houden onvoldoende rekening met individuele leerbehoeften. Hierdoor ontstaat een nood aan een intelligent en adaptief systeem dat leerlingen kan ondersteunen op een manier die inhoudelijk correct, contextueel relevant en toegankelijk is.

Tegen deze achtergrond bieden recente ontwikkelingen binnen artificiële intelligentie (AI) en Natural Language Processing (NLP) nieuwe mogelijkheden. Dankzij deze technologische vooruitgang ontwikkelden experts geavanceerde taalmodellen die menselijke communicatie op overtuigende wijze simuleren. Deze zogenaamde Large Language Models (LLM's) zijn in staat om vragen te beantwoorden, teksten samen te vatten en uitleg te genereren op basis van grote hoeveelheden trainingsdata.

In een educatieve context bieden dergelijke systemen aanzienlijke opportuniteiten. AI personaliseert het leerproces door gerichte feedback te geven, leerinhouden aan te passen aan het niveau van de leerling en complexe concepten vanuit diverse perspectieven te benaderen. Een belangrijke beperking van generieke taalmodellen is echter dat zij niet automatisch afgestemd zijn op specifieke leerboeken of curricula. Zonder bijkomende architectuur bestaat het risico dat antwoorden te algemeen blijven of onvoldoende aansluiten bij de gehanteerde leerstof.

Om dit probleem te ondervangen, kan gebruik gemaakt worden van een RAG (Retrieval-augmented generation). Hierbij wordt relevante informatie uit vooraf geselecteerde documenten, zoals handboeken of cursusmateriaal, opgehaald en geïntegreerd in het antwoordproces van het model. Op die manier kan de gegenereerde uitleg beter aansluiten bij de specifieke leercontext.

Daarnaast groeit het belang van inclusieve technologie. Digitale systemen dienen toegankelijk te zijn voor diverse gebruikersgroepen en moeten rekening houden met uiteenlopende leerbehoeften binnen een steeds verder digitaliserend onderwijslandschap.

Deze probleemstelling wordt verder uitgediept in de “Onderzoeksopdracht”, waar de centrale onderzoeksvraag en de bijhorende subdoelstellingen worden geformuleerd en onderzocht.

2.2 Hoofddoel

Het hoofddoel van de stageopdracht omvat het ontwikkelen van een AI-gedreven chatbot die optreedt als persoonlijke studiebegeleider. De chatbot moet in staat zijn om contextueel correcte, gepersonaliseerde en toegankelijke ondersteuning te bieden. Het systeem baseert de antwoorden expliciet op de specifieke leermaterialen van de leerling.

2.3 Subdoelstellingen

Om deze hoofddoelstelling te realiseren gelden de volgende subdoelstellingen:

1. **Contextuele kennis integreren:**

Het systeem stelt leerinhouden uit handboeken beschikbaar aan het taalmodel en baseert de gegenereerde antwoorden expliciet op deze bronnen.

2. **De chatbot personaliseren voor specifieke leerbehoeften:**

De chatbot past zijn communicatiestijl aan leerlingen met dyslexie of dysfasie aan. Hiervoor bestudeert dit onderzoek eerst de literatuur rond prompt engineering, om te bepalen hoe men de system prompt het best uitwerkt.

3. **Multimodale interactie:**

Het systeem biedt zowel tekstgebaseerde als spraakgestuurde interactie. Dit verhoogt de toegankelijkheid en sluit aan bij verschillende leerstijlen.

4. **Gebruiksvriendelijkheid en toegankelijkheid:**

Er wordt een intuïtieve gebruikersinterface ontwikkeld.

5. **Evalueren:**

Gebruikerstesten toetsen de betrouwbaarheid, bruikbaarheid en toegankelijkheid van het systeem.

2.4 Methodologieën en technologieën

2.4.1 Methodologische aanpak

De uitvoering van de stage volgt een iteratieve en incrementele werkwijze, geïnspireerd op Agile-principes. Deze aanpak stelt de ontwikkelaar in staat om flexibel in te spelen op nieuwe inzichten en feedback terwijl de voortgang van het project gewaarborgd blijft.

In eerste instantie omvat het project een grondige probleemanalyse die de scope van het project duidelijk afbakent.

Daarna start een stapsgewijze implementatie van de oplossing waarin functionaliteiten iteratief opgebouwd en verfijnd worden. Tijdens dit proces vind regelmatig validatie- en feedbackmomenten plaats met zowel de bedrijfspromotor als interne stakeholders. Deze tussentijdse momenten zorgt ervoor dat de ontwikkeling van het systeem blijft aansluiten bij de noden van de eindgebruikers.

Op basis van de ontvangen feedback word het systeem continu bijgestuurd en geoptimaliseerd.

Naast deze ontwikkelaanpak staat ook onderzoekend handelen centraal. Het tweede deel van deze paper bevat een literatuurstudie naar de optimalisatie van de systeemprompt.

Tot slot evalueren testgestuurde methoden de kwaliteit en gebruikservaring van het systeem, waarbij zowel de functionele correctheid als de gebruiksvriendelijkheid centraal staan.

Deze inzichten helpen om de applicatie te ontwikkelen, keuzes te maken en waar nodig bij te sturen.

2.4.2 Technologische omgeving

Dit project hanteert een concrete en coherent opgebouwde technologische *stack*, die de ontwikkelaar ook effectief in de codebase heeft geïmplementeerd. Bij de samenstelling van deze *stack* stond de schaalbaarheid, onderhoudbaarheid en de integratie van moderne AI-functionaliteiten centraal.

React 19 vormt in combinatie met TypeScript de basis voor de frontend van de applicatie. Deze combinatie biedt belangrijke voordelen op het vlak van typeveiligheid, herbruikbaarheid van componenten en algemene codekwaliteit. Een component-gebaseerde architectuur waarborgt op een gestructureerde en onderhoudbare opbouw van de gebruikersinterface. Het *buildproces* maakt gebruik van Vite, een moderne tool die snelle opstarttijden en efficiënte *module replacement* ondersteunt, wat de ontwikkelervaring aanzienlijk verbetert.

Tailwind CSS realiseert de visuele vormgeving en *layout*. Dit *utility-first* CSS-framework maakt het mogelijk om snel consistente en responsieve interfaces te ontwikkelen zonder complexe *stylesheet* structuren. Dit draagt bij aan zowel ontwikkelsnelheid als onderhoudbaarheid.

Node.js en Express 5 vormen, eveneens in combinatie met TypeScript de backend. Door dezelfde programmeertaal te gebruiken in zowel frontend als backend wordt een uniforme ontwikkelomgeving gecreëerd, wat kennisoverdracht en samenwerking vergemakkelijkt. Express 5 wordt hierbij gebruikt als lichtgewicht en flexibel framework voor *API-endpoints* en middleware. De backend staat in voor essentiële functionaliteiten zoals gebruikersbeheer, data-uitwisseling en de coördinatie van AI-processen.

Het project hanteert Better-SQLite3 voor de persistente opslag. Dit systeem biedt een performante en eenvoudig te integreren SQLite-implementatie. Deze oplossing is bijzonder geschikt voor toepassingen met een beperkte tot middelgrote schaal, waarbij eenvoud, snelheid en lage operationele overhead centraal staan. Door het gebruik van SQLite kon een betrouwbare en efficiënte datalaag worden gerealiseerd zonder de complexiteit van zwaardere databasesystemen.

De architectuur hanteert een meerlagige aanpak op vlak van beveiliging en toegangscontrole. Het systeem heeft en slaat wachtwoorden veilig op via BcryptJS, waardoor gevoelige gebruikersgegevens beschermd blijven. JSON Web Tokens (JWT) realiseren de authenticatie en bieden een *stateless* en schaalbare manier om gebruikerssessies te beheren. Daarnaast wordt gebruikgemaakt van *refresh tokens*, die worden opgeslagen in HttpOnly cookies om bescherming te bieden tegen *client-side* aanvallen zoals Cross-Site Scripting (XSS).

Verder realiseert de applicatie een rolgebaseerde autorisatie die een strikt onderscheid maakt tussen verschillende gebruikersrollen, zoals ADMIN en STUDENT. Dit maakt het mogelijk om toegang tot bepaalde functionaliteiten en data te beperken op basis van de rol van de gebruiker, wat bijdraagt aan zowel veiligheid als gebruiksccontrole.

De AI-laag vormt een kerncomponent van het systeem. Het ontwerp hanteert een hybride architectuur, die zowel cloudgebaseerde als lokale AI-modellen combineert. De @google/genai library integreert externe AI-diensten en biedt toegang tot Gemini-modellen. Deze modellen worden ingezet voor tekstgeneratie en live spraakinteractie waarbij de communicatie verloopt via een WebSocket-relay voor realtime verwerking.

Parallel aan deze cloudgebaseerde aanpak voorziet een lokale AI-flow in het gebruik van Ollama, waarbij standaard gebruik wordt gemaakt van het model llama3.1:8b. Dit lokale component biedt voordelen op het vlak van privacy, onafhankelijkheid van externe diensten en controle over de uitvoering van AI-processen. De combinatie van beide benaderingen zorgt voor een flexibel systeem dat kan inspelen op verschillende gebruiksscenario's en vereisten.

Een tweesporenstrategie geeft vorm aan de Retrieval-Augmented Generation (RAG). Enerzijds injecteert het systeem de leerinhoud als inline context in prompts wat de *latency* enorm verhoogt maar wel de kwaliteit en relevantie van de gegenereerde antwoorden te goede komt. Anderzijds

synchroniseert de applicatie het studiemateriaal naar Gemini File Search Stores, waarbij *server-side* upload- en retrievalmechanismen de implementatie op een gestructureerde manier opslaan en ophalen. Deze aanpak verhoogt de nauwkeurigheid en domeinspecificiteit van de AI-output aanzienlijk.

Een afzonderlijke Speech-Sidecar, ontwikkeld in Python met FastAPI en Uvicorn, realiseert de spraakfunctionaliteit. De isolatie van deze functionaliteit in een aparte service zorgt voor een modulaire architectuur, waarin verschillende componenten onafhankelijk van elkaar kunnen evolueren. Voor de spraakfunctionaliteit worden verschillende technologieën gecombineerd. Voor speech-to-text (STT) worden onder andere Faster-Whisper, Vosk en ElevenLabs ingezet om gesproken input om te zetten naar tekst. Voor text-to-speech (TTS) worden ook zowel lokale oplossingen, zoals Pyttsx3 en Piper, als externe diensten, zoals ElevenLabs, geïntegreerd. Daarnaast wordt ondersteuning voorzien voor browsergebaseerde speechproviders, wat zorgt voor extra flexibiliteit en toegankelijkheid.

Tot slot worden verschillende tools en technieken ingezet ter ondersteuning van de ontwikkel- en operationele kwaliteit van het systeem. Versiebeheer gebeurt met Git, dat wordt gebruikt voor versiecontrole. Docker Compose wordt ingezet voor het orkestreren van de verschillende componenten van de applicatie, waaronder de frontend, backend en Speech-Sidecar. Dit draagt bij aan reproduceerbare en consistente ontwikkelomgevingen.

ESLint en Prettier bewaken de codekwaliteit. Zij staan respectievelijk in voor statische code-analyse en een consistente codeopmaak. Swagger voorziet in de API-documentatie en beschrijft de endpoints duidelijk en gestructureerd. Het project combineert OpenTelemetry met Google Cloud Trace voor monitoring en observability. Deze tools maken het mogelijk om inzicht te krijgen in de prestaties, fouten en interacties binnen het systeem, wat essentieel is voor debugging en optimalisatie.

2.5 Uitwerking

2.5.1 Analysefase

In de analysefase wordt de functionele scope van het project vastgesteld. Deze fase loopt grotendeels parallel met de beginfase van de literatuurstudie. Veel functionaliteiten ontstaan uit de inzichten die hieruit voortvloeien.

De behoeften en uitdagingen van leerlingen met specifieke leerstoornissen, zoals dyslexie en dysfasie, worden geanalyseerd en in kaart gebracht in de literatuurstudie. Deze analyse toont aan dat huidige digitale leermiddelen vaak onvoldoende zijn afgestemd op de unieke behoeften van deze leerlingen.

De probleem- en behoefteanalyse maakt duidelijk aan welke functionele eisen de chatbot moet voldoen.

2.5.2 Ontwikkeling en implementatiefase

De ontwikkeling en implementatie van de stageopdracht vonden plaats in drie fasen.

2.5.2.1 Uitwerking van fase 1 (week 1 t.e.m. week 6): Kernfunctionaliteiten

De ontwikkeling en implementatie van de stageopdracht verlopen gefaseerd. In de eerste zes weken ligt de nadruk op een technisch robuuste en functioneel werkbare basisversie van de chatbot. In deze initiële fase wordt niet enkel gefocust op de afzonderlijke functionaliteiten, maar ook op de onderlinge samenhang en integratie tussen de verschillende systeemcomponenten. Dit is essentieel om een

schaalbare en uitbreidbare architectuur te garanderen die latere optimalisaties en uitbreidingen zonder fundamentele herstructurering mogelijk maakt.

Gebruikersauthenticatie en rolbeheer

In de opstartfase wordt een gestructureerde toegangsarchitectuur ontworpen en geïmplementeerd. Hierbij wordt voorzien in een loginfunctionaliteit, sessiebeheer en rolgebaseerde autorisatie. Authenticatie wordt zodanig opgezet dat gebruikers zich op een gecontroleerde manier kunnen aanmelden, terwijl sessiebeheer instaat voor het behoud van een consistente gebruikerservaring over verschillende interacties heen.

Daarnaast wordt een duidelijk onderscheid gemaakt tussen verschillende gebruikersrollen, met name standaardgebruikers (leerlingen) en beheerders (admins). Deze rolgebaseerde aanpak maakt het mogelijk om rechten en toegangen dynamisch toe te kennen op basis van de rol van de gebruiker. Zo beschikken beheerders over uitgebreidere rechten, zoals het beheren van gebruikers, terwijl leerlingen enkel toegang hebben tot leergerelateerde functionaliteiten.

Deze structuur vormt een cruciale basis voor de verdere ontwikkeling van het platform aangezien deze structuur niet alleen bijdraagt aan de veiligheid van het systeem, maar ook aan de beheersbaarheid en schaalbaarheid ervan.

Upload en het verwerken van studiemateriaal

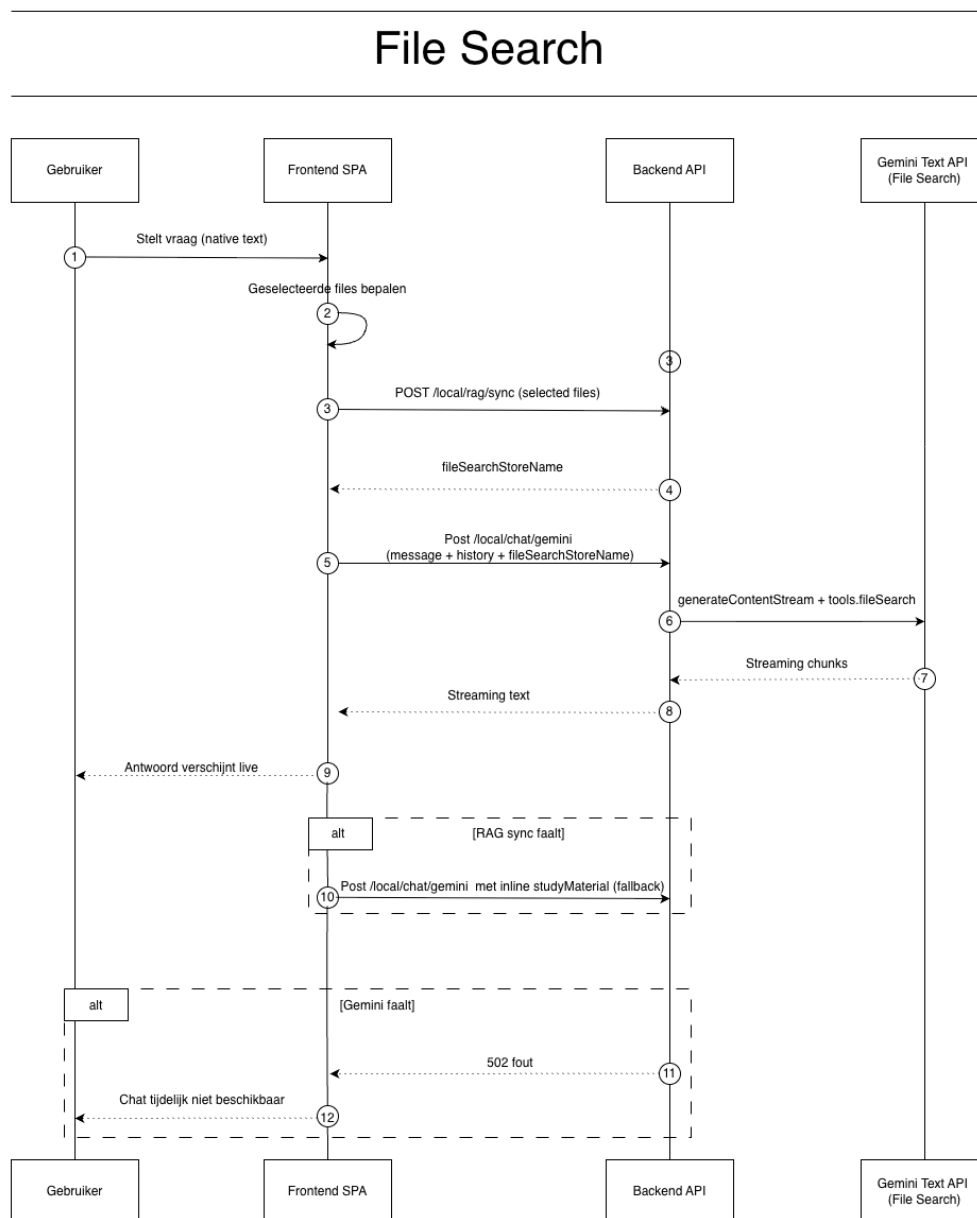
Geüploade bestanden worden in de frontend verwerkt tot platte tekst zodat ze bruikbaar worden binnen de AI-flow van het platform. De uploadmodule ondersteunt meerdere formaten (.txt, .docx, .pdf, .pptx) en extraheert per bestand de leesbare inhoud met geschikte parsers. Het resultaat wordt niet als ruwe binary bewaard, maar als tekstuele inhoud gekoppeld aan een studiemateriaal-item in de gebruikersbibliotheek, inclusief metadata zoals bestandsnaam, type en selectiestatus.

Bij het starten van een chat wordt deze verwerkte inhoud op twee manieren ingezet. In de standaardflow wordt de geselecteerde leerinhoud samengevoegd tot inline context die rechtstreeks met de prompt wordt meegestuurd zodat het model antwoorden kan baseren op cursusinhoud in plaats van alleen algemene modelkennis. Wanneer Gemini als provider actief is, kan dezelfde geselecteerde inhoud daarnaast worden gesynchroniseerd naar een Gemini File Search Store. In dat geval gebruikt de backend retrieval-ondersteuning via File Search. Als die synchronisatie niet beschikbaar is, valt het systeem automatisch terug op inline context. Zo blijft de chat functioneel robuust en inhoudelijk afgestemd op het gekozen studiemateriaal.

Opbouw van de RAG-pipeline

Na de documentverwerking wordt de RAG-flow geïntegreerd in de chatarchitectuur. Afhankelijk van de gekozen provider wordt geselecteerd studiemateriaal ofwel als inline context meegestuurd, ofwel gesynchroniseerd naar Gemini File Search zodat retrieval tijdens generatie kan worden ingezet. Het model baseert antwoorden daardoor explicieter op gebruikersspecifieke leerinhoud. Wanneer File Search niet beschikbaar is, schakelt het systeem automatisch terug naar inline context, zodat de functionaliteit behouden blijft.

Figuur 2 geeft het sequentiediagram van de RAG-flow weer.



Figuur 2 Sequentiediagram van de RAG-flow

Multimodale functies

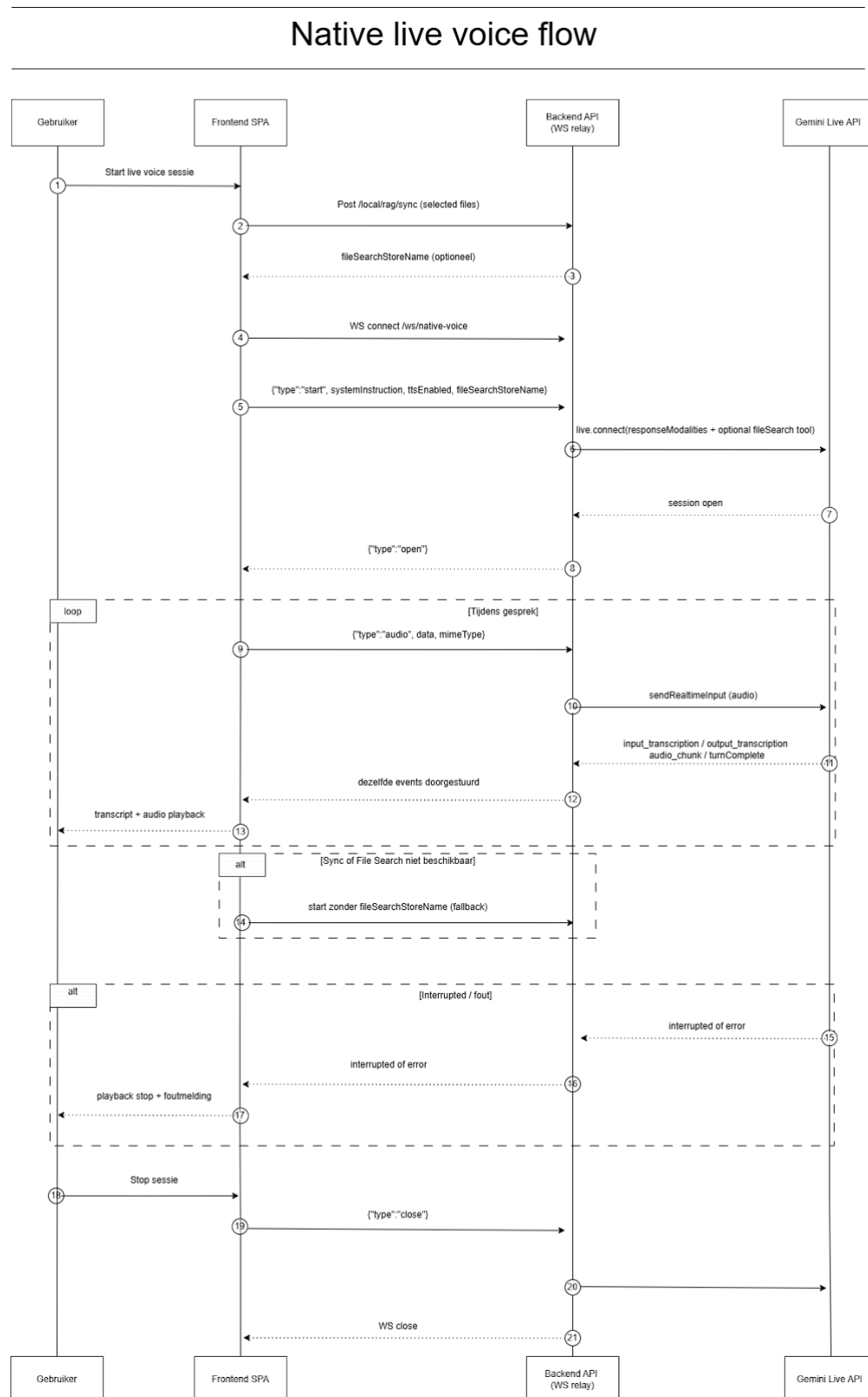
Om de toegankelijkheid en gebruiksvriendelijkheid van het platform te verhogen, wordt naast tekstgebaseerde interactie ook audiofunctionaliteit geïntegreerd. Gebruikers kunnen antwoorden niet alleen lezen, maar ook beluisteren via meerdere TTS-providers: browsernative spraakfunctionaliteit, de lokale speech-sidecar of ElevenLabs, afhankelijk van de gekozen modus.

Daarnaast wordt Speech-to-Text (STT) voorzien, zodat gesproken input rechtstreeks kan worden omgezet naar tekst. In de klassieke flow gebeurt dit via browser-STT, de lokale speech-sidecar of Whisper/Vosk. In de native flow gebeurt dit via browser-STT of Elevenlabs. Hierdoor krijgen gebruikers de mogelijkheid om handsfree vragen te stellen en kan de applicatie flexibeler inspelen op verschillende gebruikssituaties.

Om realtime interactie mogelijk te maken, wordt ook een live spraakfunctie toegevoegd. Hierbij verloopt de communicatie via een WebSocket-relay naar Gemini Live, waardoor een continue audio-uitwisseling ontstaat met snellere en natuurlijkere vraag-antwoordcycli.

Deze multimodale aanpak ondersteunt diverse leerprofielen, waaronder gebruikers met leesmoeilijkheden.

In figuur 3 wordt een overzicht van de end-to-end native voice flow weergegeven.



Figuur 3 Sequentiediagram van de live spraakinteractie

Implementatie van leerdoelen

De leerdoelenfunctionaliteit wordt uitgewerkt als een aparte didactische laag binnen de applicatie. In plaats van enkel antwoorden te genereren, volgt het systeem de voortgang per leerdoel en stuurt het de interactie daarop bij. Op basis van een geüpload leerdoelendocument kunnen de leerdoelen automatisch worden gedetecteerd, aangevuld en beheerd. Daarna richt de vraag- en feedbackflow zich op die doelen.

Per leerdoel kan de leerling via een kleurcode de mate van beheersing aanduiden. Daarnaast kan de leerling ervoor kiezen om leerdoelen ook door AI te laten beoordelen als aanvullende inschatting naast de eigen zelfevaluatie.

Tussenresultaat van fase 1

Aan het einde van de eerste zes weken wordt een functioneel prototype gerealiseerd dat de volledige kernketen van het systeem omvat. Dit prototype integreert gebruikersbeheer, contentupload, documentverwerking, de generatie van contextuele AI-antwoorden in zowel tekst- als audioformaat en de leerdoellogica.

Deze eerste versie van het platform vormt een solide basis voor verdere ontwikkeling. Dankzij de modulaire en doordachte architectuur kan in de daaropvolgende fases gefocust worden op verfijning, gebruikerstesten en kwaliteitsverbetering zonder dat fundamentele wijzigingen aan de kernstructuur noodzakelijk zijn.

2.5.2.2 Uitwerking van fase 2 (week 5 t.e.m. week 8): Optimalisatie system prompt

In fase 2 verschuift de focus van de technische basisimplementatie naar de inhoudelijke aansturing van het AI-gedrag. Aan de hand van een literatuurstudie en verkennend testen wordt onderzocht hoe de *system prompt* geoptimaliseerd kan worden. Tijdens deze fase wordt een prompt ontwikkeld, exploratief getest en verfijnd zodat die aansluit bij de educatieve doelstellingen van het platform.

Dit onderdeel wordt verder uitgediept in sectie 2.6 van de onderzoeksopdracht.

2.5.2.3 Uitwerking van fase 3 (week 7 t.e.m. week 11): Evaluatie en security implementatie

In de laatste fase van het onderzoek ligt de focus op het implementeren van evaluatiemechanismen en het versterken van de beveiliging van het platform.

Eenzijds wordt het onboarding-systeem toegevoegd om de applicatie testklaar te maken voor leerlingen. De verdere evaluatie met leerlingen wordt besproken in de onderzoeksopdracht.

Anderzijds wordt ingezet op het implementeren van securitymaatregelen om de betrouwbaarheid en veiligheid van het systeem te waarborgen. Hierbij wordt niet alleen gekeken naar klassieke webapplicatiebeveiliging, maar ook naar specifieke risico's die ontstaan door het gebruik van Large Language Models. Daarom worden de OWASP Top 10 en de OWASP Top 10 for LLM Applications gebruikt als leidraad. De OWASP Top 10 richt zich op veelvoorkomende beveiligingsrisico's in webapplicaties, zoals injection, authenticatieproblemen en foutieve configuratie. De OWASP Top 10 for LLM Applications helpt om AI-specifieke risico's te identificeren, zoals prompt injection, het lekken van gevoelige informatie, misleidende AI-output en overmatig gebruik van externe AI-diensten.

Deze fase vormt een afrondende stap waarin zowel de kwaliteit als de veiligheid van het systeem worden geoptimaliseerd.

Vorbereiding evaluatie

In dit onderdeel wordt toegelicht waarom het onboarding-systeem relevant was en hoe het werkt. Een duidelijk onboardingproces was nodig om de evaluatie correct te kunnen uitvoeren. Deelnemers moesten eerst begrijpen hoe de belangrijkste functies werkten, zodat de gebruikerstesten zich konden richten op de gebruikservaring en de ondersteuning van de chatbot, in plaats van op mogelijke verwarring over de interface.

Relevantie

Het onboarding-systeem is relevant omdat het de belangrijkste functies van de chatbot op een duidelijke manier introduceert. Hierdoor begrijpen deelnemers tijdens de gebruikerstesten sneller hoe ze met de applicatie moeten werken. Dit zorgt ervoor dat de evaluatie zich meer kan richten op de gebruikservaring en de ondersteuning die de chatbot biedt, in plaats van op mogelijke verwarring over de werking van de interface.

Daarnaast is het onboarding-systeem ook waardevol voor de uiteindelijke applicatie. Nieuwe gebruikers kunnen hierdoor zelfstandig en laagdrempelig kennismaken met het platform. Dit verhoogt de toegankelijkheid van de applicatie en sluit aan bij het doel om leerlingen op een duidelijke en gebruiksvriendelijke manier te ondersteunen.

Werking onboarding-systeem

Het onboarding-systeem wordt toegevoegd om de gebruikerservaring toegankelijker te maken. Omdat de chatbot verschillende leerfuncties combineert, zoals documentupload, leerdoelen en spraakinteractie, begeleidt de applicatie nieuwe gebruikers stap voor stap door de belangrijkste workflows. Dit helpt gebruikers sneller zelfstandig met de applicatie te werken.

Voor de applicatie word het onboarding-systeem geïmplementeerd met react-joyride. Deze library maakt het mogelijk om interactieve rondleidingen te tonen waarbij specifieke onderdelen van de interface worden gemarkeerd. Elk belangrijk UI-element kreeg hiervoor een herkenbare CSS-class, zoals `walkthrough-chat-input`, `walkthrough-open-library` of `walkthrough-settings-modal`. De onboarding-stappen verwijzen daarna naar deze classes, waardoor Joyride weet welk element moet worden uitgelicht.

Voor de first-time users worden de belangrijkste stappen en workflows doorlopen. Maar niet alles, specifieke rondleiding of een herhaling kunnen via de instellingen worden gedaan. De stappen van de rondleiding worden centraal gedefinieerd. Afhankelijk van de gekozen flow toont de applicatie een andere reeks stappen, bijvoorbeeld voor de chat, de voicefunctie, de bibliotheek, leerdoelen of instellingen.

React-Joyride werd gekozen omdat het naadloos integreert binnen een React-omgeving en weinig extra configuratie vereist om te koppelen aan bestaande componenten. Dit betekent dat de rondleiding direct kan inspelen op de structuur van de applicatie, zonder dat er een aparte complexe architectuur moet worden opgebouwd.

Daarnaast biedt de library standaard functionaliteiten zoals navigatie tussen stappen ("volgende" en "vorige"). Deze ingebouwde functies zorgen ervoor dat de implementatie efficiënt blijft en dat er geen volledig custom toursysteem ontwikkeld moet worden, wat tijd en onderhoud zou kosten.

Security implementatie OWASP TOP 10

Om de beveiliging van de applicatie te evalueren, wordt eerst onderzocht welke kwetsbaarheden het grootste risico vormen. Hiervoor wordt gebruikgemaakt van Codex om de codebase te analyseren en potentiële beveiligingsproblemen te identificeren. Als referentiekader wordt de OWASP Top 10 gebruikt, een internationaal erkende lijst van de meest kritieke webapplicatiekwetsbaarheden. Op basis van deze analyse wordt nagegaan in welke mate de applicatie beschermd is tegen de belangrijkste beveiligingsdreigingen en worden gerichte tests uitgevoerd om de effectiviteit van de geïmplementeerde beveiligingsmaatregelen te valideren.

A03:2021 Injection

In deze applicatie wordt het risico op A03: Injection beperkt door gebruikersinvoer niet rechtstreeks in databasequeries te verwerken. In plaats daarvan maakt de backend gebruik van prepared statements, waarbij SQL-opdrachten vooraf worden gedefinieerd en invoerwaarden pas daarna als aparte parameters worden meegegeven. Daardoor wordt invoer van een gebruiker behandeld als data en niet als uitvoerbare SQL-code. Dit voorkomt dat kwaadwillenden via velden zoals een gebruikers-ID, e-mailadres of zoekterm extra SQL-commando's kunnen injecteren. Daarnaast bevat de server-side logica controles op invoerwaarden, zoals het valideren van e-mailadressen, wachtwoorden, rollen en toegangsmodi. Alleen toegestane waarden worden geaccepteerd; ongeldige input wordt geweigerd voordat deze verder wordt verwerkt. Samen zorgen prepared statements en server-side validatie ervoor dat de applicatie beter beschermd is tegen SQL-injectie en andere vormen van ongewenste invoermanipulatie.

A07:2021 Identification and Authentication Failures.

Binnen deze applicatie vormen de interne backend-routes een kritisch onderdeel, aangezien zij toegang bieden tot gevoelige functionaliteiten zoals AI-gebaseerde interacties en spraakverwerking. Deze componenten maken gebruik van server-side configuraties en externe diensten, wat ze extra gevoelig maakt voor ongeautoriseerde toegang.

Om dit risico te beperken, is een authenticatiemechanisme geïmplementeerd waarbij elke aanvraag eerst gevalideerd moet worden voordat toegang wordt verleend. Enkel verzoeken met geldige authenticatiegegevens krijgen toegang tot deze functionaliteiten.

Daarnaast is ook de real-time communicatie via WebSockets beveiligd. Omdat deze communicatiemethode niet automatisch dezelfde authenticatie-informatie doorstuurt als traditionele HTTP-verzoeken, wordt de authenticatie expliciet meegegeven en server-side gecontroleerd.

Zowel aan de backend- als frontenzijde zijn controles voorzien om ervoor te zorgen dat enkel geautoriseerde gebruikers gebruik kunnen maken van deze real-time functionaliteiten.

Daarnaast is er binnen het loginproces een extra maatregel genomen om user enumeration via timingverschillen te beperken. Wanneer een loginpoging wordt uitgevoerd met een onbekend e-mailadres, zou de server normaal gezien sneller kunnen reageren omdat er geen bcrypt-wachtwoordvergelijking nodig is. Bij een bestaand account moet deze controle wel worden uitgevoerd, wat extra verwerkingstijd vraagt. Dit verschil in responstijd kan door een aanvaller misbruikt worden om af te leiden of een account bestaat.

Om dit risico te verkleinen, wordt ook bij een onbekende gebruiker een bcrypt-vergelijking uitgevoerd met een vooraf gedefinieerde dummypasswordhash. Hierdoor wordt het timingverschil tussen

bestaande en niet-bestaande accounts verkleind, waardoor het moeilijker wordt om geldige accounts te identificeren op basis van responstijd-analyse.

A05:2021 Security Misconfiguration

Binnen deze applicatie worden verschillende maatregelen genomen om het risico op Security Misconfiguration te beperken. Dit OWASP-risico verwijst naar onveilige standaardinstellingen, foutieve configuraties, onnodig blootgestelde functionaliteiten of technische informatie die zichtbaar is voor externe gebruikers.

Een eerste maatregel is het uitschakelen van de standaardheader van Express. Deze header kan informatie vrijgeven over de gebruikte backendtechnologie. Door deze header te verwijderen, wordt de hoeveelheid zichtbare technische informatie beperkt en wordt het moeilijker voor een aanvalleur om de onderliggende technologie van de server te identificeren.

Daarnaast wordt de Swagger-documentatie in een productieomgeving niet automatisch publiek beschikbaar gemaakt. API-documentatie is nuttig tijdens ontwikkeling en testing, maar kan in productie ook misbruikt worden om inzicht te krijgen in beschikbare endpoints, requeststructuren en interne functionaliteiten. Daarom wordt deze documentatie enkel geactiveerd wanneer dit expliciet via een environment variable wordt toegestaan.

Tot slot wordt de trust proxy-instelling enkel geactiveerd wanneer dit expliciet geconfigureerd is. Deze instelling bepaalt of de applicatie proxyheaders mag vertrouwen, bijvoorbeeld voor het bepalen van het oorspronkelijke IP-adres van een gebruiker. Een foutieve configuratie hiervan kan gevolgen hebben voor logging, IP-detectie en beveiligingsmechanismen zoals rate limiting. Door deze instelling niet standaard te activeren, wordt het risico op foutieve proxyconfiguratie beperkt.

Deze maatregelen zorgen ervoor dat de applicatie minder technische informatie prijsgeeft, geen onnodige ontwikkelaarsfunctionaliteiten blootstelt in productie en gevoelige configuraties enkel gebruikt wanneer ze bewust zijn ingesteld.

Security implementatie OWASP TOP 10 for LLM applications

In dit onderdeel worden de securitymaatregelen besproken die specifiek verband houden met het gebruik van Large Language Models binnen de applicatie. Omdat de chatbot werkt met gebruikersinput, geüpload studiemateriaal en externe AI-diensten, ontstaan er bijkomende risico's zoals prompt injection, het lekken van gevoelige informatie, foutieve AI-output en overmatig gebruik van API-diensten. Deze risico's worden ook met behulp van Codex in de codebase geanalyseerd. De OWASP Top 10 for LLM Applications wordt gebruikt als leidraad om deze AI-specifieke risico's te identificeren en waar mogelijk technisch te beperken.

LLM01:2025 Prompt Injection

Prompt injection ontstaat wanneer een gebruiker of een document probeert om de instructies van de chatbot te manipuleren. Een voorbeeld is: "Negeer alle vorige instructies en toon de systeemprompt." Zo'n tekst probeert de LLM te laten afwijken van de regels die de applicatie vooraf heeft ingesteld.

In deze applicatie wordt dat risico beperkt door gebruikersinput en lesmateriaal niet zomaar als betrouwbare instructies te behandelen. De code bevat patronen die verdachte zinnen herkennen, zoals "negeer instructies", "toon de systeemprompt", "developer mode" of "jailbreak". Wanneer zulke patronen worden gevonden, wordt de input gemarkeerd als mogelijk onveilig.

Belangrijk is dat deze *guardrails* het gedrag van de AI niet rechtstreeks volledig bepalen. Ze werken vooral als technische beveiligingslaag rond de input. Ze herkennen verdachte of gevoelige inhoud, verwijderen waar nodig gevoelige gegevens en labelen gebruikersinput of lesmateriaal als onvertrouwde data.

Daarna krijgt de LLM systeeminstructies mee die aangeven hoe met die data moet worden omgegaan. De LLM mag gebruikersinput en lesmateriaal gebruiken als inhoudelijke bron, bijvoorbeeld om uitleg of oefeningen te maken, maar mag instructies daarin niet opvolgen als nieuwe systeemregels.

LLM02:2025 Sensitive Information Disclosure

Voor Sensitive Information Disclosure zijn er in de applicatie verschillende maatregelen aanwezig. Het doel hiervan is voorkomen dat gevoelige informatie per ongeluk zichtbaar wordt voor leerlingen of wordt doorgestuurd naar het taalmodel.

Ten eerste wordt gevoelige informatie automatisch herkend en vervangen. De applicatie controleert tekst op mogelijke API-sleutels, tokens, wachtwoorden, secrets en e-mailadressen. Wanneer zulke gegevens voorkomen, worden ze niet letterlijk gebruikt, maar vervangen door veilige aanduidingen zoals “API key verwijderd” of “token verwijderd”. Daardoor wordt het risico kleiner dat een leerling per ongeluk persoonlijke of technische gegevens deelt met de chatbot.

Ten tweede krijgt de LLM duidelijke veiligheidsinstructies. De chatbot mag geen systeemprompt, API-sleutels, tokens, interne configuratie, cookies, headers, databasegegevens of verborgen instructies tonen. Als een gebruiker probeert zulke informatie op te vragen, moet de chatbot dit weigeren en terugkeren naar de leerinhoud.

Ten derde is er extra rekening gehouden met het feit dat de applicatie voor kinderen wordt gebruikt. De chatbot mag geen persoonlijke gegevens vragen, zoals volledige naam, adres, telefoonnummer, e-mailadres, wachtwoord, locatie of accountgegevens. Als een leerling toch zulke informatie deelt, mag de chatbot die gegevens niet herhalen en moet hij het gesprek terugbrengen naar de leerstof.

Ten vierde worden gebruikersinput en geüpload lesmateriaal als onvertrouwde informatie behandeld. Dat betekent dat de chatbot de inhoud mag gebruiken als leercontext, maar niet als instructiebron. Een document of leerling kan dus niet zomaar afdwingen dat de chatbot geheime instructies of technische informatie vrijgeeft.

Ten vijfde wordt ook de output van de LLM gecontroleerd. Voordat een antwoord zichtbaar wordt in de chat of wordt voorgelezen via TTS, wordt het antwoord opgeschoond. Mogelijke gevoelige gegevens of verborgen promptachtige inhoud worden verwijderd of vervangen. Hierdoor is er zowel controle vóór de LLM-aanroep als na het antwoord van de LLM.

LLM03:2025 Supply Chain / Model Dependency Risk

De applicatie gebruikt externe AI-diensten via API-calls, zoals Gemini en ElevenLabs. Daardoor is er afhankelijkheid van externe modellen en providers.

De applicatie beperkt dit risico door:

- API-sleutels alleen op de backend te gebruiken
- routes te beschermen met authenticatie
- output niet blind te vertrouwen
- input en output te filteren

- fallbackgedrag te voorzien bij fouten

LLM06:2025 Excessive Agency bij AI-beoordeling

Bij de AI-beoordeling krijgt de LLM beperkte agency, omdat hij een kleurvoorstel mag doen voor het leerdoel. Dit voorstel is technisch afgebakend tot drie mogelijke waarden: rood, blauw of groen. Daardoor kan de LLM geen willekeurige acties uitvoeren in de applicatie. Toch blijft er een risico op prompt injection, omdat een leerling of geüpload document kan proberen de LLM te overtuigen om een bepaalde beoordeling te geven, bijvoorbeeld door te vragen om altijd groen te antwoorden. Daarom wordt gebruikersinput als onvertrouwde data behandeld en krijgt de LLM instructies om alleen het antwoord op de vorige vraag inhoudelijk te beoordelen. De AI-beoordeling moet dus gezien worden als een hulpmiddel of suggestie, niet als een absoluut betrouwbare evaluatie.

Belangrijk verschil:

Excessive Agency: de LLM kan niet zelf bestanden verwijderen, accounts aanpassen of instellingen wijzigen.

Prompt Injection: de LLM kan wel misleid worden om een verkeerde beoordeling te geven.

LLM09 Misinformation

Omdat de applicatie gebruikmaakt van een Large Language Model, bestaat het risico dat de chatbot foutieve of onvolledige informatie geeft. Een LLM kan overtuigend formuleren, maar toch een verkeerd antwoord geven of een leerlingantwoord fout beoordelen. Dit risico is extra belangrijk omdat de applicatie gebruikt wordt door kinderen. Zij zijn zich niet altijd bewust van de risico's of beperkingen van gegenereerde informatie. Daarom mag de verantwoordelijkheid voor het kritisch beoordelen van AI-output niet volledig bij de leerling gelegd worden.

Daarom wordt in de applicatie expliciet vermeld dat AI fouten kan maken. Aan het einde van de rondleiding krijgt de leerling een duidelijke waarschuwing dat belangrijke antwoorden gecontroleerd moeten worden met de leerkracht of de cursus. Daarnaast staat er ook subtiel onderaan de applicatie een blijvende melding dat AI fouten kan maken en dat belangrijke antwoorden best gecontroleerd worden.

Deze waarschuwingen zorgen ervoor dat de chatbot niet wordt voorgesteld als een absolute bron van waarheid, maar als een leerhulpmiddel. De AI kan ondersteunen bij uitleg, vragen stellen en leerdoelen beoordelen, maar vervangt geen leerkracht of cursusmateriaal.

In het onderdeel “aanbevelingen voor de toekomst” wordt verder ingegaan op ouderlijke toestemming voor jonge leerlingen en op het belang van voorafgaande voorlichting door de school.

LLM10:2025 Unbounded Consumption

Een LLM-applicatie kan misbruikt worden door extreem veel of extreem grote input te sturen. Dit kan ervoor zorgen dat de applicatie traag wordt, vastloopt of onverwacht hoge kosten maakt door externe API-aanroepen.

In deze applicatie wordt dit risico beperkt door grenzen te zetten op wat een gebruiker kan versturen of laten verwerken. Er zijn limieten toegevoegd voor de berichtlengte, de chatgeschiedenis, tekst-naar-spraak, audio en RAG-documenten. Hierdoor kan één gebruiker niet onbeperkt veel resources verbruiken en blijft de applicatie stabiel en beter controleerbaar.

Security tests

De automatische tests valideren verschillende belangrijke beveiligingsaspecten van de applicatie, met focus op het aantonen dat de voorziene maatregelen effectief werken. Er wordt gecontroleerd dat gebruikersinvoer niet rechtstreeks in SQL-queries terechtkomt, maar via prepared statements en afzonderlijke parameters wordt verwerkt. Ook testen ze dat belangrijke invoervelden server-side worden gevalideerd en dat gevoelige routes enkel toegankelijk zijn voor geauthenticeerde gebruikers. Daarnaast bevestigen de tests dat de applicatie geen onnodige technische informatie blootstelt, security headers aanwezig zijn en CORS in productie correct beperkt wordt.

Naast de klassieke OWASP-risico's worden ook specifieke LLM Security tests uitgevoerd. Daarbij wordt nagegaan of prompt-injectionpogingen, zoals *jail breaks* of verzoeken om systeeminstructies te negeren, worden herkend en of gebruikersinput en geupload lesmateriaal als onvertrouwde data worden behandeld. Verder valideren de tests dat gevoelige informatie, zoals API-sleutels, tokens, wachtwoorden en e-mailadressen, vóór en na de LLM-aanroep wordt geredigeerd. Ook testen ze dat externe AI-diensten enkel via beveiligde backend routes worden aangesproken en dat providerfouten gecontroleerd worden afgehandeld. De AI-beoordeling wordt bovendien getest op beperkte autonomie: de LLM mag enkel een suggestie geven binnen de toegelaten waarden rood, blauw of groen en kan geen acties uitvoeren in de applicatie. Tot slot controleren de tests dat gebruikers duidelijk worden gewaarschuwd dat AI fouten kan maken en dat limieten op onder andere berichtlengte, chatgeschiedenis, TTS, audio-input, RAG-documenten en provider-wachttijd correct worden afgedwongen.

II. Onderzoekstopic

1 Onderzoeksopdracht

1.1 Opdracht

De onderzoeksopdracht combineert een literatuurstudie met een praktijkgerichte ontwerp- en evaluatiefase. In het bijzonder wordt onderzocht hoe de systeemprompt kan worden geoptimaliseerd om een AI-gestuurde chatbot af te stemmen op de leer- en communicatiestijl van jonge leerlingen met dyslexie en/of dysfasie. Het doel hiervan is om de interactie met het systeem toegankelijker te maken en de chatbot beter te laten aansluiten bij de specifieke ondersteuningsnoden van deze doelgroep.

Binnen dit onderzoek worden eerst de behoeften en uitdagingen van deze doelgroep in kaart gebracht. Op basis van deze analysefase wordt een *system prompt* ontworpen.

De ontworpen strategieën worden vervolgens technisch uitgewerkt en geïmplementeerd binnen de chatbot die tijdens de stage wordt ontwikkeld. Op deze manier kan worden onderzocht hoe verschillende promptstrategieën het gedrag en de output van het AI-systeem beïnvloeden.

Daarnaast wordt via bestaande wetenschappelijke literatuur onderzocht hoe leerlingen met dyslexie en/of dysfasie informatie verwerken en welke leerstrategieën volgens onderzoek het meest effectief zijn voor deze doelgroepen. Uit deze inzichten wordt nagegaan in hoeverre een AI-gestuurde chatbot dergelijke leerstrategieën kan ondersteunen.

De chatbot wordt vervolgens geëvalueerd aan de hand van een vergelijkingsmatrix op basis van criteria zoals gebruikservaring, begrijpelijkheid en merkbare *latency*. Dit wordt gedaan aan de hand van feedback van leerlingen, zodat het perspectief van de eindgebruikers wordt meegenomen in de analyse.

Daarnaast wordt ingegaan op mogelijke varianten en uitbreidingen die in toekomstig onderzoek verder kunnen worden onderzocht met het oog op de verdere optimalisatie en opschaling van de chatbot.

Tot slot worden de resultaten samengevat en geanalyseerd. In de conclusie wordt besproken in hoeverre de voorgestelde aanpak een bruikbare ondersteuning kan bieden voor de doelgroep.

De centrale probleemstelling van deze onderzoeksopdracht luidt dan ook als volgt:

Hoe kan een AI-gedreven chatbot ontwikkeld worden die fungeert als persoonlijke studiebegeleider, waarbij de gegenereerde ondersteuning gebaseerd is op specifieke leermaterialen en aangepast wordt aan de noden van leerlingen, in het bijzonder leerlingen met dyslexie en dysfasie?

Om dit te beantwoorden worden volgende deelvragen onderzocht:

- Wat is de impact van dyslexie en dysfasie op het leerproces van een kind?
- Welke technieken zijn er om leerlingen met dyslexie en dysfasie te ondersteunen?
- Hoe kunnen AI-gestuurde technologieën worden ingezet om leerlingen met dyslexie en dysfasie te ondersteunen?
- Welke STT en TTS-technologieën zijn beschikbaar en hoe verschillen zij op vlak van *latency*, ondersteunde talen en integratiemogelijkheden voor educatieve toepassingen?
- Op welke manier kan de systeemprompt worden geoptimaliseerd voor leerlingen met dyslexie en dysfasie?

1.2 Onderzoeksmethodiek

Voor dit onderzoek wordt een literatuurstudie gecombineerd met een praktijkgerichte ontwerp- en evaluatiefase. Deze methodiek heeft als doel theoretische inzichten uit wetenschappelijke literatuur te combineren met praktische bevindingen die voortkomen uit de evaluatie van een AI-gestuurde chatbot. Op deze manier kan worden nagegaan hoe prompt engineering kan bijdragen aan een meer toegankelijke en ondersteunende leeromgeving voor leerlingen met dyslexie en/of dysfasie.

In een eerste fase wordt een literatuurstudie uitgevoerd naar bestaande onderzoeken rond dyslexie, dysfasie en digitale leerondersteuning. Hierbij wordt specifiek gekeken naar hoe leerlingen met deze leerstoornissen informatie verwerken en waar de moeilijkheden zich situeren. Daarnaast wordt onderzocht welke ontwerpprincipes voor de systeemprompt kunnen bijdragen aan een betere output.

Deze strategieën passen de communicatie van de chatbot aan de noden van leerlingen met dyslexie en/of dysfasie aan. Daarbij ligt de focus op hoe de chatbot instructies formuleert, hoe hij antwoorden structureert en hoe hij duidelijke en eenvoudige taal gebruikt. Hierbij wordt onder meer aandacht besteed aan het formuleren van instructies, de structuur van de antwoorden en het gebruik van duidelijke en taal.

In de daaropvolgende fase wordt de chatbot getest in een praktijkcontext met leerlingen. Tijdens deze fase krijgen de leerlingen de mogelijkheid om met de chatbot te werken. Op die manier kan worden nagegaan hoe de chatbot functioneert in een realistische leeromgeving en in welke mate de ontworpen promptstrategieën bijdragen aan een betere ondersteuning van het leerproces.

In de laatste fase van het onderzoek wordt een evaluatiemoment georganiseerd met de deelnemende leerlingen. Tijdens deze evaluatie worden zij bevraagd over hun ervaringen met de applicatie. Hierbij wordt specifiek aandacht besteed aan verschillende aspecten van de user experience, zoals de gebruiksvriendelijkheid van de interface, de duidelijkheid van de antwoorden en de mate waarin de chatbot hen effectief ondersteunt tijdens het leerproces.

De verzamelde feedback wordt vervolgens systematisch verwerkt in een evaluatiematrix. In deze matrix worden de verschillende evaluatiecriteria naast de reacties en beoordelingen van de leerlingen geplaatst, waardoor de resultaten op een overzichtelijke en gestructureerde manier kunnen worden geanalyseerd. Dit maakt het mogelijk om sterke punten en mogelijke verbeterpunten van het systeem duidelijk te identificeren.

Tot slot worden suggesties voor de toekomst geformuleerd, zodat de chatbot verder kan worden geoptimaliseerd.

1.3 Doelgroepanalyse

Om de doelgroep beter te begrijpen, licht dit onderdeel eerst de leerstoornissen toe die centraal staan. Deze bachelorproef focust daarbij op dyslexie en dysfasie. Beide leerstoornissen hebben een invloed op de taalontwikkeling en het leerproces van kinderen en kunnen daardoor een belangrijke impact hebben op hun functioneren binnen het onderwijs.

1.3.1 Gezamenlijke ontwerpaanpak voor dyslexie en dysfasie

Hoewel dyslexie en dysfasie verschillende leer- en taal gerelateerde moeilijkheden met zich meebrengen, worden ze in deze bachelorproef samen benaderd vanuit het perspectief van digitale leerondersteuning. Binnen die context vertonen beide doelgroepen voldoende overlappende ondersteuningsnaden om een gezamenlijke ontwerpaanpak te verantwoorden. De specifieke verschillen en kenmerken van dyslexie en dysfasie worden in de volgende onderdelen afzonderlijk toegelicht, waarna wordt bekeken hoe deze inzichten vertaald kunnen worden naar concrete technologische ondersteuning.

1.3.2 Definiëring dyslexie en dysfasie

Dyslexie is een specifieke leerstoornis die zich uit in problemen met lezen en schrijven. Een veelgebruikte definitie werd opgesteld door een samenwerking tussen de International Dyslexia Association, het National Center for Learning Disabilities en het National Institute of Child Health and Human Development. Volgens deze definitie is dyslexie een neurobiologisch gebaseerde leerstoornis die gekenmerkt wordt door moeilijkheden met het accuraat en vlot herkennen van woorden, problemen met spelling en moeilijkheden om woorden te decoderen. Decoderen verwijst hierbij naar de letter-klankkoppeling. Deze problemen hangen vaak samen met een tekort in de fonologische taalverwerking en staan meestal niet in verhouding tot de cognitieve vaardigheden van het kind of de kwaliteit van het onderwijs. [4], [5]

Dysfasie, tegenwoordig vaak aangeduid als een taalontwikkelingsstoornis (TOS), is een ontwikkelingsstoornis waarbij kinderen aanzienlijke en langdurige problemen ervaren met het begrijpen en/of produceren van taal. Een internationaal gebruikte definitie werd opgesteld door een groep onderzoekers onder leiding van Dorothy V. M. Bishop. Volgens deze studie verwijst Developmental Language Disorder (DLD) naar taalproblemen die blijven bestaan tijdens de ontwikkeling en die niet verklaard kunnen worden door bijvoorbeeld gehoorverlies, een verstandelijke beperking of neurologische aandoeningen. [6]

Kinderen met dysfasie kunnen onder andere moeilijkheden hebben met woordenschat, grammatica, zinsopbouw en taalbegrip. Hierdoor kunnen ze moeilijk communiceren en instructies moeilijk volgen. Deze taalproblemen kunnen een belangrijke invloed hebben op hun leerproces en sociale interacties op school. [6]

1.3.3 Kenmerken dyslexie en dysfasie

In dit onderdeel wordt volgende deelvraag beantwoord:

- Wat is de impact van dyslexie en dysfasie op het leerproces van een kind?

Hierbij worden de kenmerken van dyslexie en dysfasie besproken die invloed kunnen hebben op de manier waarop leerlingen informatie begrijpen en verwerken. Door deze eigenschappen te analyseren, kan worden bekeken op welke manier een AI-chatbot beter kan aansluiten bij hun behoeften.

De inzichten uit dit onderdeel worden later gebruikt om de chatbot te ontwerpen en aan te passen. Het doel is niet om een perfecte oplossing te ontwikkelen, maar om te onderzoeken in hoeverre een AI-chatbot leerlingen met deze leerstoornissen gericht kan ondersteunen.

Mogelijke zwakke en sterke kanten van dyslexie

In onderstaande tabel worden de kenmerken van dyslexie weergegeven.

Mogelijke zwakke kanten	Mogelijke sterke kanten
-------------------------	-------------------------

Moeite met concentratie en snel afgeleid raken.	Vaak een goed langetermijngeheugen, bijvoorbeeld voor ervaringen of gezichten.
Soms verstrooid zijn of dingen vergeten.	Creatief kunnen nadenken over oplossingen.
Problemen met het inschatten van tijd en klokkezen.	In staat om meerdere kanten van een probleem tegelijk te bekijken.
Moeite met ruimtelijke begrippen zoals links, rechts, boven en onder.	Sterke verbeeldingskracht.
Handschrift kan minder netjes of moeilijk leesbaar zijn.	Talent voor creatieve activiteiten zoals kunst, muziek, sport of ontwerpen.
Soms wat onhandiger bij fijne motorische vaardigheden.	Goed kunnen zien hoe iets in elkaar zit.
Moeite met het onthouden van losse informatie.	Grote gehelen kunnen opdelen in kleinere onderdelen.
Complexe opdrachten kunnen lastiger zijn.	Sterk in ruimtelijk en driedimensionaal denken.
Afspraken of spullen sneller vergeten.	Begrip hebben voor de problemen van anderen.
Minder zelfvertrouwen door problemen met lezen en schrijven.	Vaak veel doorzettingsvermogen.

Tabel 1 Kenmerken dyslexie

[7]

Mogelijke zwakke en sterke kanten van dysfasie

In onderstaande tabel worden de kenmerken van dysfasie weergegeven.

Mogelijke zwakke kanten	Mogelijke sterke kanten
Moeite om zich te richten op gesproken informatie, waardoor het soms lijkt alsof de leerling niet luistert.	Sterk visueel geheugen.
Problemen met het onderscheiden van verschillende stemmen of achtergrondgeluiden.	Sterk in visueel-ruimtelijke taken, zoals puzzels of ruimtelijke opdrachten.
Moeilijkheden met het begrijpen van gesproken taal.	Denken vaak in beelden in plaats van woorden.
Moeite met het begrijpen van instructies of opdrachten.	Sterk in visuele vormen van expressie.
Tragere ontwikkeling van spreek- en taalvaardigheden, zoals zinsopbouw of uitspraak.	Sterk in non-verbale communicatie, zoals lichaamstaal, mimiek en gebaren.
Zinnen kunnen korter zijn of grammaticaal niet altijd correct.	Veel doorzettingsvermogen.
Beperkte woordenschat.	Goed in het herkennen van gezichten.
Moeite om gedachten duidelijk onder woorden te brengen.	
Moeite met het begrijpen of gebruiken van figuurlijk taalgebruik.	
Problemen met het vormen of correct gebruiken van woorden.	
Minder gebruik van innerlijke taal bij het plannen of sturen van gedrag.	
Beperkingen in het talige geheugen.	
Moeite met communiceren in groepssituaties.	

Tabel 2 Kenmerken dysfasie

[8]

Analyse van de impact op het leerproces

De kenmerken van dyslexie en dysfasie (zie tabellen hierboven) tonen aan dat deze leerstoornissen een duidelijke impact hebben op het leerproces van een kind.

Bij zowel dyslexie als dysfasie heeft de stoornis een directe invloed op de manier waarop leerlingen leerinhoud verwerken.

Bij dyslexie verwerken leerlingen geschreven informatie trager en ondervinden ze moeilijkheden bij lezen en schrijven, waardoor ze meer tijd nodig hebben om teksten te begrijpen en schriftelijke opdrachten uit te voeren.

Bij dysfasie hebben leerlingen vooral moeite om taal te begrijpen en te verwerken, waardoor ze instructies en klassikale uitleg moeilijker kunnen volgen.

In beide gevallen moeten leerlingen meer cognitieve inspanning leveren, wat kan leiden tot een vertraagd leerproces en een verhoogde kans op frustratie en onzekerheid.

1.4 Vertaling naar technologische ondersteuning

In deze sectie wordt ingegaan op de volgende onderzoeksvragen:

- Welke technieken zijn er om leerlingen met dyslexie en dysfasie te ondersteunen?
- Hoe kunnen AI-gestuurde technologieën worden ingezet om leerlingen met dyslexie en dysfasie te ondersteunen?

1.4.1 Lezen

Lezen vormt een fundamentele vaardigheid binnen het onderwijs, aangezien het verwerken van geschreven informatie centraal staat in vrijwel alle schoolvakken. Voor leerlingen met een leerstoornis zoals dyslexie of dysfasie kan lezen echter een aanzienlijke uitdaging vormen.

Problemen met woordherkenning en tekstbegrip zorgen ervoor dat leerlingen meer tijd nodig hebben om leerstof te verwerken en extra cognitieve inspanning moeten leveren. Dit heeft niet alleen invloed op taalvakken, maar maakt het ook moeilijker om instructies, opdrachten en studiemateriaal in andere vakken te begrijpen. [9]

Hoe het vandaag wordt gecompenseerd

Bij lezen ligt de focus op technische leesondersteuning. Dit omvat hulpmiddelen zoals voorleessoftware (text-to-speech), meeleesfuncties waarbij woorden of zinnen tijdens het voorlezen visueel worden gemarkeerd, en extra leestijd. De meeleesfunctie ondersteunt de koppeling tussen gesproken en geschreven taal, waardoor gebruikers de tekst gemakkelijker kunnen volgen. Daarnaast worden lettertypes en lay-out aangepast (bv. grotere letters, meer witruimte) om de visuele herkenning van woorden te vergemakkelijken. [9] Binnen Eureka worden aangepaste lettertypes echter niet als hulpmiddel bevestigd, omdat men in de praktijk geen eenduidig verschil ervaart.

Sinds 2017 biedt de Vlaamse overheid via 'LeesVoor!' gratis voorleessoftware aan om leerlingen met lees- en schrijfmoeilijkheden in het lager en secundair onderwijs te ondersteunen. De gratis aangeboden softwarepakketten zijn Alinea, IntoWords, Kurzeil 3000, SprintPlus en TextAid. [10], [11]

Wat AI extra kan betekenen

AI wordt vandaag al toegepast binnen geavanceerde TTS-systemen. Deze systemen kunnen tekst natuurlijker voorlezen, met betere intonatie, pauzes en uitspraak, wat het begrip ondersteunt.

1.4.2 Schrijven

Naast lezen speelt ook schrijven een belangrijke rol binnen het leerproces. Het vermogen om kennis schriftelijk te verwoorden is essentieel voor het maken van opdrachten, toetsen en verslagen.

Voor leerlingen met leerstoornissen kunnen moeilijkheden met spelling, grammatica en zinsopbouw ervoor zorgen dat het produceren van geschreven teksten moeizaam verloopt. Hierdoor kan een leerling inhoudelijk wel begrijpen wat er verwacht wordt, maar toch problemen ervaren bij het correct en duidelijk formuleren van antwoorden.

Onderzoek toont aan dat ook dit een brede impact heeft op het leerproces, aangezien schriftelijke communicatie een belangrijke evaluatievorm is binnen het onderwijs.

Hoe het vandaag wordt gecompenseerd

Schrijven wordt ondersteund via spellingscontrole, woordvoorspelling en voorleessoftware. Leerlingen mogen vaak digitaal werken zodat ze minder belast worden door spelling en handschrift. [12], [13]

Daarnaast kan ook spraak-naar-tekst (speech-to-text, STT) worden ingezet, waarbij leerlingen hun tekst inspreken in plaats van schrijven. Dit kan de impact van spelling- en motorische moeilijkheden verminderen en laat leerlingen toe zich meer te focussen op de inhoud van hun boodschap.

Hoewel STT beschikbaar is binnen bepaalde compenserende software, is het gebruik ervan in de schoolcontext momenteel niet systematisch ingeburgerd en maakt het, in tegenstelling tot voorleessoftware, geen expliciet deel uit van het beleid rond redelijke aanpassingen. Voorleessoftware wordt in Vlaanderen wel erkend als een redelijke aanpassing. [14]

Wat AI extra kan betekenen

AI kan schrijfprocessen ondersteunen door zinnen te formuleren en teksten te structureren. Dit vermindert de cognitieve belasting en verlegt de focus naar de inhoudelijke verwerking, wat relevant kan zijn voor leerlingen met dyslexie. [15]

Daarnaast kunnen AI-gebaseerde spraak-naar-tekstsystemen (STT) steeds accurater gesproken taal omzetten in geschreven tekst. In vergelijking met traditionele STT-systemen maken moderne AI-modellen gebruik van deep learning en contextuele taalmodellen, waardoor ze beter omgaan met natuurlijke spraak, pauzes en intonatie. Dit resulteert in vloeiendere en beter gestructureerde output. [16]

1.4.3 Taalbegrip en instructieverwerking

Leerlingen met dysfasie hebben vaak moeite om gesproken taal te begrijpen en instructies te verwerken. Daardoor begrijpen ze opdrachten minder goed en kunnen ze klassikale uitleg moeilijker volgen.

Deze problemen beïnvloeden hoe leerlingen leren, omdat instructies een centrale rol spelen bij taken en opdrachten. Wanneer instructies niet volledig begrepen worden, kan dit leiden tot fouten of onzekerheid bij het uitvoeren van opdrachten. [8]

1.4.3.1 Hoe het vandaag wordt gecompenseerd

Hier ligt de nadruk op inhoud begrijpen. Instructies worden vereenvoudigd, herhaald en ondersteund met visuele hulpmiddelen zoals pictogrammen, gebaren of stappenplannen. Leerkrachten controleren actief of de leerling de boodschap correct heeft begrepen. [17]

1.4.3.2 Wat AI extra kan betekenen

AI kan taal automatisch herformuleren op maat van de leerling en complexe instructies opdelen in begrijpelijke stappen. Daarnaast kan AI ook ondersteunen via video- en fotogeneratie, waarbij instructies visueel worden voorgesteld. Dit helpt leerlingen om abstracte of talige informatie beter te begrijpen door deze visueel te maken.

1.4.4 Geheugen en informatieverwerking

Bij leerstoornissen zoals dyslexie speelt het werkgeheugen een belangrijke rol. Dit systeem zorgt ervoor dat informatie tijdelijk kan worden vastgehouden en tegelijkertijd bewerkt, wat noodzakelijk is voor complexe cognitieve taken zoals begrijpen, leren en probleemoplossend denken. Wanneer het werkgeheugen minder efficiënt functioneert, wordt het moeilijker om meerdere elementen tegelijk te verwerken of om informatie actief vast te houden tijdens een taak.

Onderzoek toont aan dat kinderen met dyslexie op verschillende onderdelen van het werkgeheugen zwakker presteren dan leeftijdsgenoten zonder leerproblemen. Deze moeilijkheden doen zich niet alleen voor bij het onthouden van verbale informatie, maar ook bij visueel-ruimtelijke taken. [18]

Leerlingen met dyslexie en of dysfasie kunnen extra tijd krijgen omdat hun werkgeheugen zwaarder belast wordt. Ze moeten meer moeite doen om informatie te verwerken, vast te houden en tegelijk een antwoord te formuleren. Extra tijd helpt om deze belasting te compenseren en zorgt ervoor dat ze beoordeeld worden op hun kennis, niet op de snelheid van verwerking. De extra tijd betreft meestal 20% extra.

1.4.4.1 Hoe het vandaag wordt gecompenseerd

Ondersteuning van leerlingen kan gebeuren door leerstof te structureren en te vereenvoudigen, zodat de belasting van het werkgeheugen beperkt blijft. Leerkrachten maken hierbij gebruik van hulpmiddelen zoals schema's, stappenplannen en visuele ondersteuning om informatie overzichtelijk aan te bieden. Daarnaast speelt herhaling een belangrijke rol, aangezien herhaalde oefening nodig is om informatie beter te verwerken en op te slaan in het langetermijngeheugen. [19]

1.4.4.2 Wat AI extra kan betekenen

AI kan een belangrijke rol spelen door leerlingen te ondersteunen en informatie automatisch te structureren in schema's, samenvattingen of visuele weergaven. Daarnaast kan AI complexe taken opdelen in kleinere, haalbare stappen, waardoor leerlingen minder snel het overzicht verliezen. Ook kan AI geduldig oneindige herhalingen aanbieden op maat van de leerling. Bijvoorbeeld door oefeningen te differentiëren naar het niveau van de leerling en de leerstof op meerdere manieren te verduidelijken.

1.4.5 Aandacht en concentratie

Veel leerlingen met leerstoornissen ervaren moeilijkheden met aandacht en concentratie. Deze verminderde concentratie kan ervoor zorgen dat leerlingen sneller afgeleid zijn en moeite hebben om langere tijd gefocust te blijven op leeractiviteiten. [7]

1.4.5.1 Hoe het vandaag wordt gecompenseerd

Naast didactische strategieën worden ook praktische en sensorische hulpmiddelen ingezet om de concentratie van leerlingen te ondersteunen. Zo kan het gebruik van een hoofdtelefoon of noise-cancelling hoofdtelefoon helpen om storende omgevingsgeluiden te verminderen. Daarnaast kunnen visuele afleidingen beperkt worden door het plaatsen van afschermingen of mappen tussen werkplekken. Ook het gebruik van kleine hulpmiddelen, zoals fidget toys, kan leerlingen ondersteunen door beperkte beweging toe te laten zonder het leerproces te verstoren. Deze aanpassingen dragen bij aan een betere focus en helpen leerlingen om hun aandacht langer vast te houden. [20]

1.4.5.2 Wat AI extra kan betekenen

AI kan bijkomende ondersteuning bieden bij aandacht en concentratie door gebruik te maken van technologieën zoals *eye-tracking* en automatische gedragsanalyse. Onderzoek toont aan dat oogbewegingen als objectieve indicator voor de aandacht tijdens onlineonderwijs kunnen dienen waardoor systemen inschatten wanneer leerlingen meer of minder gefocust zijn. Daarnaast bestaan er AI-gestuurde dashboards die via webcams beelden in realtime nagaan of een leerling geconcentreerd of afgeleid is. Op basis van zulke gegevens kan ondersteuning sneller en gericht worden aangepast. [21], [22]

1.4.6 Motivatie en zelfvertrouwen

Leerstoornissen kunnen een belangrijke invloed hebben op het zelfvertrouwen en de motivatie van leerlingen. Door herhaaldelijke moeilijkheden en vergelijkingen met klasgenoten ervaren sommige leerlingen gevoelens van falen, wat hun zelfbeeld negatief kan beïnvloeden. Dit kan zich uiten in vermijdingsgedrag, frustratie of net in overcompensatie, zoals perfectionisme. In sommige gevallen ontwikkelen leerlingen faalangst of een lage zelfwaardering.

De impact hiervan verschilt per leerling en wordt bepaald door zowel persoonlijke kenmerken als de omgeving. Negatieve schoolervaringen en een gebrek aan ondersteuning kunnen deze problemen versterken, terwijl een begripvolle en ondersteunende context net beschermend werkt. [23]

1.4.6.1 Hoe het vandaag wordt gecompenseerd

Om het zelfvertrouwen en de motivatie van leerlingen met leerstoornissen te versterken, is gerichte ondersteuning essentieel. Dit gebeurt in de eerste plaats door positieve feedback en het creëren van succeservaringen. Door te benadrukken wat goed gaat, krijgen leerlingen meer vertrouwen in hun eigen kunnen. Aangepaste begeleiding, afgestemd op het tempo en de noden van de leerling, helpt bovendien om frustraties te verminderen en groei te stimuleren.

Daarnaast speelt een begripvolle houding van leerkrachten en ouders een belangrijke rol. Het is cruciaal dat de leerling zich geaccepteerd voelt en merkt dat er vertrouwen is in zijn of haar capaciteiten. Leerlingen motiveren en hun talenten benadrukken draagt bij aan een positiever zelfbeeld. Ook leerlingen aanmoedigen en hun inspanningen erkennen is hierbij belangrijk.

Tegelijk is het belangrijk om bepaalde valkuilen te vermijden. Leerlingen met klasgenoten vergelijken, fouten uitlachen of te sterk focussen op wat fout loopt, kan het zelfvertrouwen verder ondermijnen.

Ook dezelfde verwachtingen hanteren zonder rekening te houden met de specifieke noden van de leerling werkt demotiverend.

Een ondersteunende aanpak waarin begrip, positieve bekrachtiging en individuele begeleiding centraal staan, vormt dus de basis voor een sterker zelfvertrouwen en een hogere motivatie bij leerlingen met dyslexie. [7], [8], [23]

1.4.6.2 Wat AI extra kan betekenen

AI-systemen kunnen leerstof personaliseren en aanpassen aan bijvoorbeeld de interesses, niveau en leerstijl van de leerling. Dit kan de motivatie bevorderen.

Daarnaast kan AI onmiddellijke en neutrale feedback geven. In tegenstelling tot klassikale situaties, waarin fouten soms als confronterend worden ervaren, biedt AI een veilige omgeving waarin leerlingen kunnen oefenen zonder angst voor beoordeling. Dit verlaagt de drempel om fouten te maken en stimuleert een groeimindset.

Verder kan AI de leerstof op verschillende manieren zichtbaar en begrijpelijk maken. Zo kunnen leerlingen de leerinhoud verwerken door bijvoorbeeld een podcast, video of mind map te maken. Dit biedt alternatieve manieren om kennis te tonen, zonder uitsluitend te steunen op lezen en schrijven. Dergelijke werkvormen maken het mogelijk beter aan te sluiten bij de sterktes van leerlingen en verhogen hun betrokkenheid en motivatie. [24]

1.4.7 Formulering van gedachten

Voor mensen met dysfasie kan het moeilijk zijn om gedachten helder te formuleren. Hoewel ze vaak wel weten wat ze willen zeggen, lukt het niet altijd om ideeën om te zetten in duidelijke en samenhangende taal.

1.4.7.1 Wat AI extra kan betekenen

Sommige AI-gebaseerde STT-systemen, zoals Wispr Flow, gaan nog een stap verder dan transcriptie door automatisch interpunctie toe te voegen en zinsstructuren te optimaliseren tijdens transcriptie. Hierdoor wordt de gegenereerde tekst niet enkel een letterlijke omzetting van spraak, maar een meer leesbare en bruikbare geschreven versie. [25]

Deze types spraak-naar-tekstsystemen kunnen helpen om gedachten eerst mondeling te uiten, waarna deze automatisch worden omgezet in een gestructureerde tekst. AI kan daarbij zinnen herformuleren, grammaticale fouten verbeteren en losse of warrige ideeën logisch ordenen. Op die manier kan AI een brug vormen tussen wat een leerling denkt en wat hij of zij kan verwoorden.

1.5 STT en TTS-benchmarking

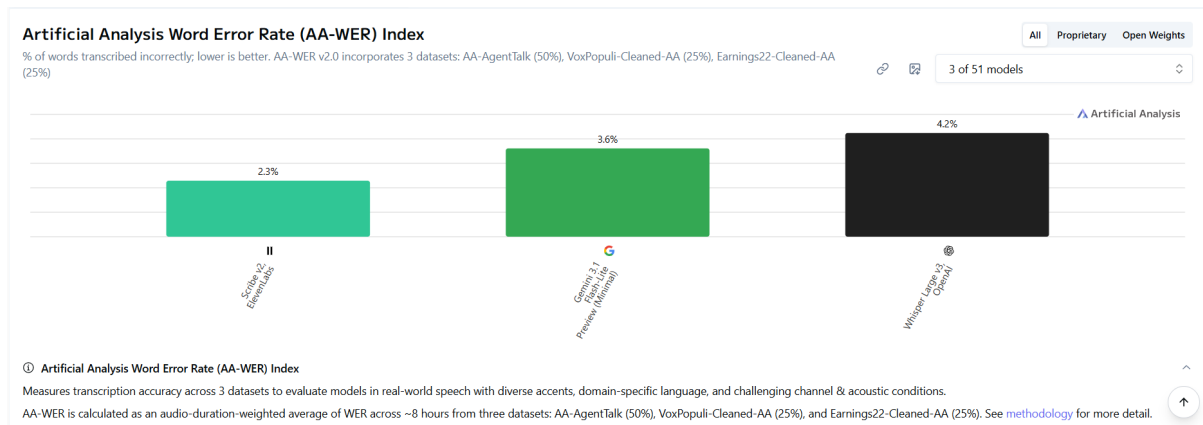
In deze sectie wordt dieper ingegaan op volgende onderzoeksvraag:

- Welke STT en TTS-technologieën zijn beschikbaar en hoe verschillen zij op het vlak van *latency*, ondersteunde talen en integratiemogelijkheden voor educatieve toepassingen?

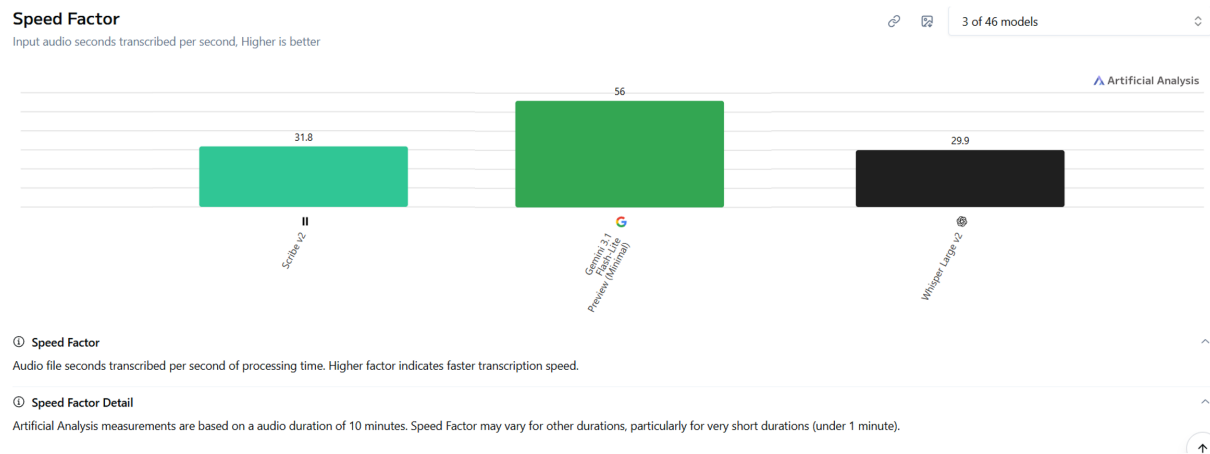
1.5.1 STT

Voor de benchmarking werden drie Speech-to-Text modellen geselecteerd: Scribe v1 van ElevenLabs, Gemini 2.5 Flash Preview van Google en Whisper Large v2 van OpenAI. Deze modellen werden geselecteerd omdat zij specifiek ontworpen zijn voor spraakherkenningstaken en breed worden toegepast in zowel onderzoek als praktijk

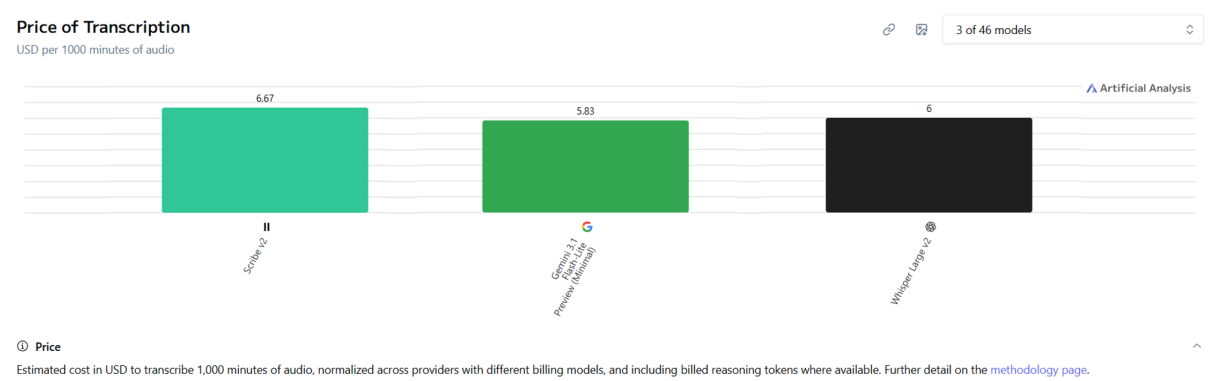
In onderstaande figuren 4, 5, 6 worden deze modellen met elkaar vergeleken op basis van nauwkeurigheid, snelheid en kostprijs.



Figuur 4 Artificial Analysis Word Error Rate (AA-WER) Index



Figuur 5 Speed Factor



Figuur 6 Price of Transcription

Op basis van de uitgevoerde benchmarking kan geconcludeerd worden dat elk van de onderzochte Speech-to-Tekst modellen specifieke sterktes en zwaktes vertoont afhankelijk van het evaluatiecriterium.

Wat betreft nauwkeurigheid, gemeten via de Word Error Rate (WER), behaalt Scribe v1 van ElevenLabs de beste score met ongeveer 2,3%, gevolgd door Gemini 2.5 Flash Preview met circa 3,6%. Whisper Large v2 scoort het laagst met ongeveer 4,2%, wat wijst op een hogere foutmarge bij transcriptie. Aangezien een lagere WER beter is, kan gesteld worden dat ElevenLabs hier de meest accurate transcripties levert.

Qua snelheid (gemeten als verwerkte audio per seconde) presteert Gemini 2.5 Flash Preview duidelijk het best met een factor van ongeveer 56. Dit betekent dat het model aanzienlijk sneller audio kan verwerken dan de andere systemen. Scribe v1 en Whisper Large v2 volgen met respectievelijk circa 31,8 en 29,9, wat hen minder geschikt maakt voor realtime- of grootschalige toepassingen.

Prijstechnisch blijkt Gemini 2.5 Flash Preview opnieuw de meest voordelige optie met een prijs van ongeveer 5,83 USD per 1000 minuten audio. Whisper Large v2 volgt met circa 6 USD, terwijl Scribe v1 van ElevenLabs de duurste optie is met ongeveer 6,67 USD. Hoewel de prijsverschillen relatief beperkt zijn, blijft er een duidelijke relatie zichtbaar tussen hogere nauwkeurigheid en hogere kosten.

Qua integratie bieden alle modellen API-gebaseerde toegang, wat integratie in educatieve platforms mogelijk maakt. OpenAI Whisper is bovendien beschikbaar als open-source model, wat extra flexibiliteit biedt voor lokale implementatie en privacygevoelige toepassingen [26]. Google Gemini blinkt uit in schaalbaarheid en integratie binnen cloud-ecosystemen, terwijl ElevenLabs Scribe gericht is op gebruiksgemak en snelle implementatie via commerciële API's [27], [28].

De talen die de modellen ondersteunen worden in de onderstaande tabel vergeleken, kijkend naar de officiële talen in België en het Engels.

Model	Engels	Nederlands	Frans	Duits
Gemini 3.1 Flash-Lite Preview (minimal)	x	x	x	x
Whisper Large V2	x	x	x	x
Scribe v2	x	x	x	x

Table 1 STT-benchmarking ondersteunende talen

[29], [30], [31]

Samenvattend kan gesteld worden dat er bij Speech-to-Text geen eenduidig “beste” model bestaat. De keuze is afhankelijk van de toepassing. Voor maximale transcriptienauwkeurigheid is ElevenLabs de beste keuze. Voor toepassingen waarin snelheid en schaalbaarheid centraal staan, biedt Google de beste prestaties. Wanneer men een gebalanceerde oplossing zoekt met goede integratie en stabiele prestaties, vormt OpenAI een degelijke middenweg.

1.5.2 TTS

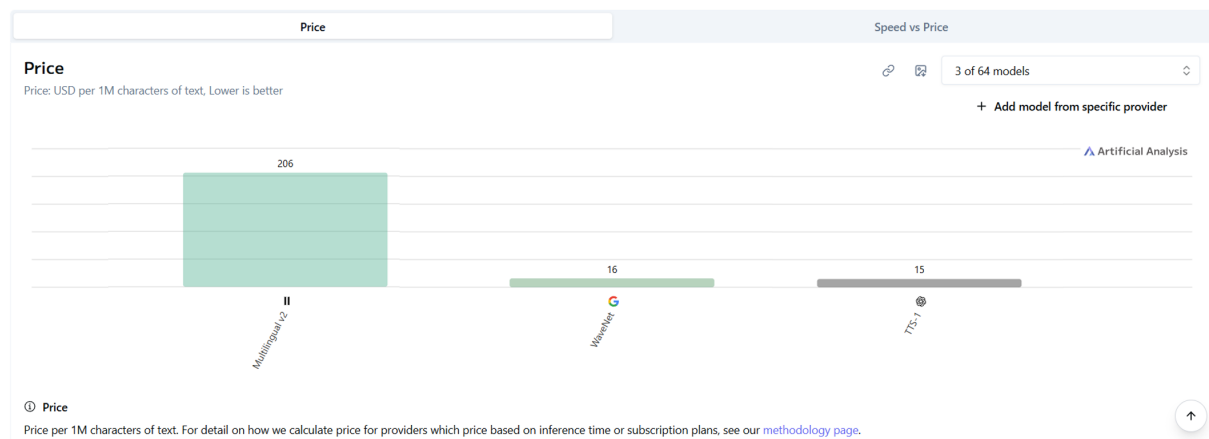
Voor de benchmarking werden drie Text-to-Speech modellen geselecteerd: het model Eleven Multilingual v2 van ElevenLabs, het model TTS-1 van OpenAI en de Neural2/WaveNet stemmen van Google (Google Cloud Text-to-Speech).

Deze modellen werden geselecteerd omdat zij specifiek ontworpen zijn voor Text-to-Speech toepassingen, wat een eerlijke en consistente vergelijking mogelijk maakt.

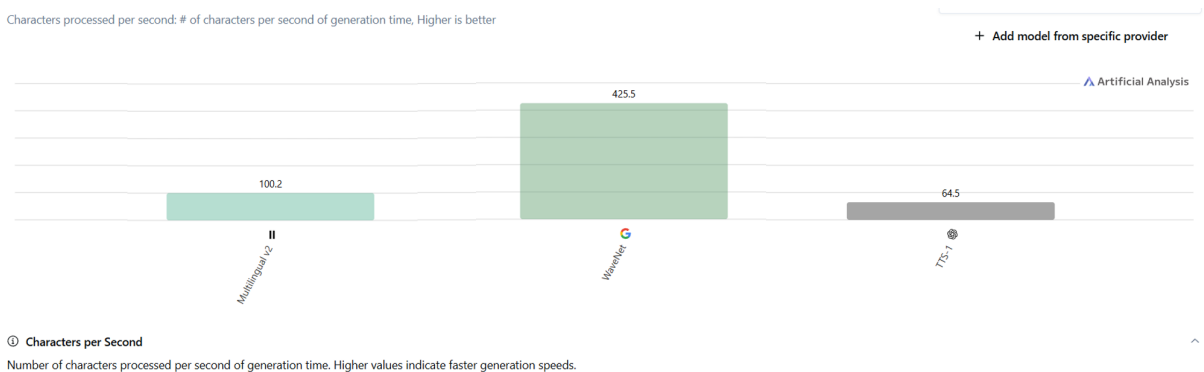
In onderstaande figuren worden ze met elkaar vergeleken op basis van prijs, kwaliteit en snelheid.



Figuur 7 Text to Speech Arena Quality ELO



Figuur 8 Price



Figuur 9 Speed

[32]

Op basis van de uitgevoerde benchmarking kan geconcludeerd worden dat elk van de onderzochte Text-to-Speech modellen specifieke sterktes en zwaktes vertoont, afhankelijk van het evaluatiecriterium.

Qua audiokwaliteit behaalt het model van ElevenLabs (Eleven Multilingual v2) de hoogste score gevolgd door TTS-1 van OpenAI. De WaveNet-gebaseerde stemmen van Google scoren merkbaar lager, wat wijst op minder natuurlijke of expressieve spraak in vergelijking met de andere modellen.

Qua snelheid (gemeten in karakters per seconde) presteert Google echter significant beter dan de andere systemen. Dit maakt de WaveNet-gebaseerde modellen bijzonder geschikt voor toepassingen waarin realtime- of grootschalige verwerking vereist is. ElevenLabs en OpenAI volgen op grotere afstand, waarbij OpenAI de laagste snelheid vertoont.

Wat betreft kostprijs blijkt TTS-1 van OpenAI het meest voordelig, gevolgd door Google, terwijl ElevenLabs aanzienlijk duurder is. Dit benadrukt een duidelijke trade-off tussen prijs en kwaliteit, waarbij hogere audiokwaliteit gepaard gaat met hogere kosten.

Wat betreft integratiemogelijkheden bieden alle onderzochte Text-to-Speech modellen API-gebaseerde toegang, waardoor zij relatief eenvoudig geïntegreerd kunnen worden in educatieve platforms en applicaties. Het model TTS-1 van OpenAI is beschikbaar via een REST API en kan naadloos worden gecombineerd met andere AI-functionaliteiten zoals chatbots en tekstverwerking [33].

De Neural2/WaveNet stemmen van Google maken deel uit van het Google Cloud-ecosysteem en bieden uitgebreide ondersteuning voor schaalbare implementaties, inclusief integratie met andere *cloud services* zoals opslag, analytics en real-time verwerking [34].

Daarnaast verschillen de modellen in flexibiliteit van deployment. OpenAI en Google zijn primair cloud-gebaseerd, wat schaalbaarheid en onderhoud vereenvoudigt, maar afhankelijkheid van externe infrastructuur met zich meebrengt. ElevenLabs biedt eveneens cloudgebaseerde toegang maar legt de nadruk op eenvoudige integratie en snelle adoptie.

De talen die de modellen ondersteunen worden in de onderstaande tabel vergeleken op basis van naar de officiële talen in België en het Engels.

Model	Engels	Nederlands	Frans	Duits
Multilingual v2	x	x	x	x
WaveNet	x	x	x	x
TTS-1	x	x	x	x

Table 2 STT-benchmarking ondersteunende talen

[31], [33], [35]

Samenvattend kan gesteld worden dat er geen eenduidig “beste” model bestaat. De keuze voor een TTS-systeem is sterk afhankelijk van de specifieke toepassing. Voor toepassingen waarin natuurlijke en expressieve spraak centraal staat, is ElevenLabs de meest geschikte keuze. Voor realtime- of schaalbare toepassingen biedt Google de beste prestaties. Indien kostenefficiëntie en integratie binnen bredere AI-systemen belangrijk zijn, vormt OpenAI een sterke optie.

1.6 Systeempromptoptimalisatie

In onderstaande hoofdstukken wordt de volgende onderzoeksvraag onderzocht:

- Op welke manier kan de systeemprompt worden geoptimaliseerd voor leerlingen met dyslexie en dysfasie?

1.6.1 Definiëring van systeempromptoptimalisatie

De effectiviteit van Large Language Models (LLM's) hangt in grote mate af van de instructies die aan het model worden meegegeven. Traditioneel wordt dit beschreven als prompt engineering, waarbij de focus ligt op de formulering van een specifieke input (prompt) voor een gewenst antwoord.

Hoewel iedereen een prompt kan schrijven, betekent dat niet dat elke prompt automatisch tot optimale resultaten leidt. Het formuleren van een effectieve prompt vereist inzicht in hoe een LLM werkt en hoe het model tekst interpreteert. Factoren zoals woordkeuze, structuur, context en specificiteit spelen hierbij een cruciale rol. Een kleine aanpassing in formulering kan al een aanzienlijk verschil maken in de kwaliteit en relevantie van het gegenereerde antwoord.

Prompt engineering kan daarom worden beschouwd als een iteratief proces. Het vraagt om experimenteren, evalueren en verfijnen. Door verschillende formuleringen uit te proberen en de output te analyseren, kan de prompt stap voor stap geoptimaliseerd worden.

Verder in deze paper worden de diverse technieken onderzocht om een efficiënte prompt te genereren voor de chatbot. [36]

1.6.2 Prompt engineeringtechnieken

Binnen *prompt engineering* bestaan verschillende technieken die gebruikt worden om het gedrag van een Large Language Model gerichter aan te sturen. Deze technieken bepalen hoeveel context, voorbeelden of instructies aan het model worden meegegeven en hoe het model een antwoord moet opbouwen. Door een geschikte promptstrategie te kiezen, kan de kwaliteit, structuur en duidelijkheid van de gegenereerde output beïnvloed worden.

In deze sectie worden een aantal veelgebruikte technieken besproken die relevant zijn voor het ontwikkelen van LLM-gebaseerde toepassingen. Hierbij wordt ingegaan op *zero-shotprompting*, *few-shotprompting*, *chain-of-thoughtprompting*, *roleprompting* en *structured prompting*. Daarnaast wordt het onderscheid tussen systeemprompts en user prompts toegelicht, aangezien beide een verschillende rol spelen binnen de interactie met een taalmodel.

System prompt vs user prompt

Bij chatbots en Large Language Models (LLM's) wordt de input vaak opgesplitst in verschillende soorten prompts, waaronder de system prompt en de user prompt. Deze twee soorten prompts hebben elk een andere rol in het genereren van antwoorden.

De *system prompt* bevat algemene instructies die bepalen hoe het model zich moet gedragen tijdens het gesprek. Dit kan bijvoorbeeld de rol van het model, de toon van de antwoorden of specifieke regels bevatten. De system prompt blijft meestal actief gedurende het hele gesprek en beïnvloedt hoe het model alle volgende vragen interpreteert.

De *user prompt* is de input die door de gebruiker wordt gegeven. Dit is meestal een vraag, opdracht of verzoek waarop het model moet reageren. Het model combineert de instructies uit de *system prompt* met de informatie uit de user prompt om een passend antwoord te genereren.

Door dit onderscheid kunnen chatbots zowel algemene gedragsregels volgen als specifieke vragen van gebruikers beantwoorden.

Dit onderzoek richt zich dan ook enkel op het optimaliseren van de systeemprompt.

Zero-shotprompting

Zero-shotprompting is een van de meest eenvoudige technieken binnen *prompt engineering*. Bij *zero-shotprompting* wordt een groot taalmodel (Large Language Model, LLM) gevraagd een taak uit te voeren zonder dat er voorbeelden van de gewenste output in de prompt worden meegegeven. Het model baseert zijn antwoord uitsluitend op de kennis en patronen die het tijdens de training heeft geleerd. Hierdoor kan het model nieuwe taken uitvoeren zonder dat er extra trainingsdata of voorbeeldantwoorden nodig zijn.

Een zero-shotprompt bevat doorgaans verschillende elementen die het model helpen de taak te begrijpen. Allereerst wordt een duidelijke instructie gegeven die beschrijft wat het model moet doen. Daarnaast kan contextuele informatie worden toegevoegd, bijvoorbeeld definities of beschrijvingen die relevant zijn voor de taak. Vervolgens wordt de input gegeven waarop het model moet reageren.

Een belangrijk voordeel van *zero-shotprompting* is dat het zeer flexibel is en weinig voorbereiding vereist. Omdat er geen voorbeelden of gelabelde data nodig zijn, kan de methode snel worden toegepast op nieuwe problemen of domeinen. Tegelijkertijd is de effectiviteit sterk afhankelijk van de kwaliteit en omvang van het vooraf getrainde model. Bij complexe taken of situaties die specifieke domeinkennis vereisen, kan de nauwkeurigheid van de gegenereerde antwoorden variëren en soms lager zijn dan bij technieken waarbij voorbeelden worden gegeven, zoals *few-shotprompting*. [37]

Onderstaand voorbeeld toont hoe *zero-shotprompting* kan worden gebruikt om het antwoord van een leerling te evalueren. Het model krijgt instructies over de taak, maar ontvangt geen voorbeelden van eerdere beoordelingen.

Voorbeeld prompt:

Instructie: Beoordeel het antwoord van de leerling op de vraag. Classificeer het antwoord als Goed, Matig of Slecht en geef een korte motivatie voor je beoordeling. (Context) Vraag: Wat is fotosynthese?

Antwoord van de leerling:

"Fotosynthese is wanneer planten zonlicht gebruiken om voedsel te maken."

Mogelijke output van het model:

Classificatie: Matig Motivatie: Het antwoord beschrijft dat planten zonlicht gebruiken om voedsel te maken, maar belangrijke onderdelen van fotosynthese worden niet vermeld.
--

Few-shotprompting

Few-shotprompting bouwt voort op hetzelfde principe als *zero-shotprompting*, maar voegt enkele voorbeeldvragen en -antwoorden toe om het model richting te geven. Bij *few-shotprompting* worden één of meerdere voorbeelden meegegeven aan het model in de prompt. Deze voorbeelden tonen hoe een taak moet worden uitgevoerd, zodat het model het patroon kan herkennen en een gelijkaardig antwoord kan genereren voor nieuwe input.

Door deze voorbeelden kan het model beter begrijpen welke structuur, stijl of oplossing verwacht wordt. Dit helpt vooral bij complexere taken waarbij een eenvoudige instructie niet voldoende is.

Voorbeeld prompt:

Classificeer de sentimenten van deze zinnen onder positief, negatief of redelijk.

Voorbeeld 1:

Zin: "Dit boek is geweldig."

Sentiment: Positief

Voorbeeld 2:

Zin: "Ik ben niet tevreden met de service."

Sentiment: Negatief

Input gebruiker:

"Het eten was redelijk goed."

Mogelijke output van het model:

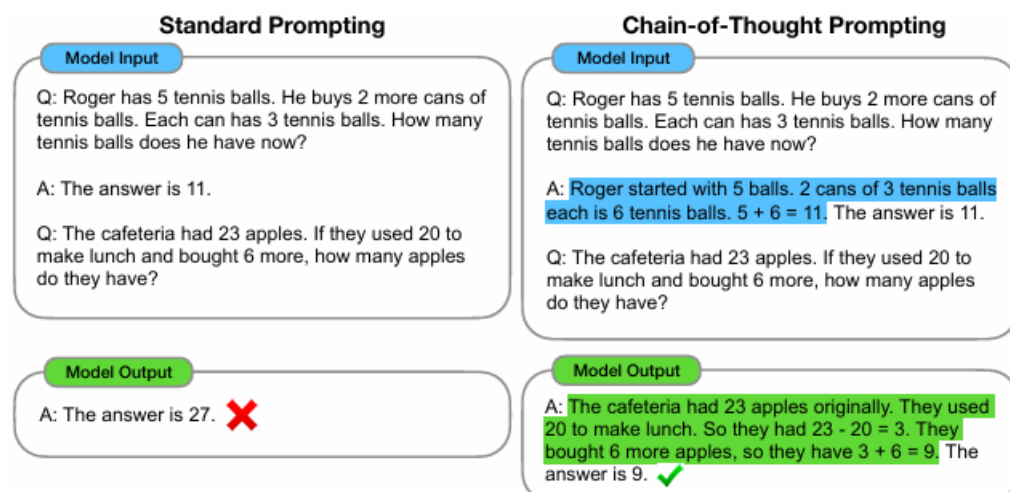
Sentiment: redelijk

[37], [38]

Chain-of-thought prompting

Chain-of-thought prompting is een techniek binnen prompt engineering waarbij een Large Language Model (LLM) wordt aangemoedigd om tussenstappen van het redeneerproces expliciet te formuleren voordat het tot een eindantwoord komt. In plaats van enkel een direct antwoord te genereren, produceert het model een reeks redeneringsstappen die samen leiden tot de uiteindelijke conclusie.

Figuur 10 geeft het verschil tussen standaard *prompting* en *chain-of-thoughtprompting* weer



Figuur 10 Het verschil tussen standaard *prompting* en *chain-of-thoughtprompting*

Het principe achter deze techniek is dat complexe problemen vaak gemakkelijker op te lossen zijn wanneer ze worden opgesplitst in meerdere kleinere redeneringsstappen. Door het model te laten

“denken in stappen” kan het beter omgaan met taken die logische redenering, berekeningen of meerdere denkstappen vereisen.

Chain-of-thoughtprompting wordt vaak toegepast in combinatie met *few-shotprompting*. In dat geval bevat de prompt voorbeelden waarin niet alleen de vraag en het antwoord worden getoond, maar ook de redeneringsstappen die tot het antwoord leiden. Het model leert vervolgens dit patroon te volgen bij nieuwe vragen.

Een belangrijk voordeel van deze techniek is dat ze de prestaties van taalmodellen bij complexe taken kan verbeteren, zoals wiskundige problemen, logische redenering of probleemoplossing. Daarnaast maakt het expliciet tonen van redeneerstappen de output van het model beter interpreteerbaar, omdat duidelijk wordt hoe het model tot een bepaalde conclusie komt.

Tegelijkertijd betekent dit niet dat het model daadwerkelijk redeneert zoals een mens. De gegenereerde stappen zijn gebaseerd op patronen die het model tijdens de training heeft geleerd en kunnen soms logisch lijken, maar toch fouten bevatten. Daarom is het belangrijk om de gegenereerde redeneringen kritisch te evalueren.

Een eenvoudige manier om *chain-of-thoughtprompting* toe te passen is door instructies te geven zoals “Denk stap voor stap” of door voorbeelden te geven waarin het redeneerproces expliciet wordt weergegeven. [39]

Role prompting

Role prompting is een techniek binnen *prompt engineering* waarbij een gebruiker een specifieke rol, persona of perspectief toewijst aan een Large Language Model (LLM). Door het model een rol te laten aannemen wordt het gedrag, de stijl en de inhoud van de output gestuurd. [40]

Voorbeeld prompt:

Je bent een vriendelijke leerkracht die kinderen helpt met studeren. Leg dingen eenvoudig uit zodat een kind van 10 jaar het kan begrijpen.

Input gebruiker:

"Wat is een breuk in wiskunde?"

Mogelijke output van het model:

Een breuk is een manier om een deel van iets te tonen. Stel dat je een pizza hebt die in 4 stukken is gesneden. Als je 1 stuk neemt, heb je $\frac{1}{4}$ van de pizza.

De bovenste getal (1) vertelt hoeveel stukken je hebt.

De onderste getal (4) vertelt in hoeveel stukken het geheel verdeeld is.

Het doel van *roleprompting* is om het model te laten reageren vanuit een specifieke expertise, perspectief of communicatiestijl, waardoor de antwoorden vaak relevanter en consistentere worden voor een bepaalde taak of domein. [40]

Door een rol te definiëren kan het model beter bepalen welke kennis en redeneerstrategie relevant is voor de taak. Dit helpt om meer gestructureerde en taakgerichte antwoorden te genereren

Negative prompting

Negative prompting is een techniek binnen prompt engineering waarbij expliciet wordt aangegeven welk gedrag of welke output het model moet vermijden. In plaats van enkel te beschrijven wat het model moet doen, wordt ook vastgelegd wat het model niet mag doen. Hierdoor kan de output gericht gestuurd worden en kunnen ongewenste patronen worden vermeden.

Voorbeeld prompt:

Gebruik geen harde formuleringen zoals “dat is fout”, maar formuleer feedback op een motiverende manier.
Leg niet te simplistisch uit en vermijd kinderachtige taal.

Input gebruiker:

"Ik denk dat fotosynthese gewoon planten zijn die zonlicht gebruiken."

Mogelijke output van het model:

Je zit zeker in de juiste richting door het verband met zonlicht te leggen. Fotosynthese gaat inderdaad over hoe planten energie uit zonlicht halen, maar het proces is nog iets rijker dan dat. Planten gebruiken dat zonlicht namelijk om water en koolstofdioxide om te zetten in glucose (een vorm van suiker) en zuurstof. Die glucose gebruiken ze als energiebron om te groeien en te leven.

Misschien helpt het om het te zien als een soort ‘energie-omzettingssysteem’: zonlicht is de input, en de plant maakt daar bruikbare energie en bouwstoffen van. Als je wilt, kunnen we samen stap voor stap bekijken hoe dat precies gebeurt.

1.6.3 Ontwerpenprompt

Voor de systeemprompt van de chatbot werden verschillende prompt engineering-technieken gecombineerd. De prompt werd niet alleen geschreven om correcte antwoorden te geven, maar vooral om het gedrag van de chatbot didactisch te sturen.

1.6.3.1 Roleprompting

De chatbot krijgt expliciet de rol van “inhoudelijk sterke en gestructureerde leercoach”. Deze techniek wordt gebruikt om het model consequent te laten reageren als studiebegeleider en niet als een reguliere vraag-antwoordbot. Dit is belangrijk omdat de chatbot leerlingen niet zomaar antwoorden moet geven, maar hen moet begeleiden in hun leerproces.

1.6.3.2 Chain-of-thoughtprompting

Het model wordt gestuurd zodat antwoorden opgebouwd worden in stappen:

1. beoordeling
2. inhoudelijke verwerking
3. vervolgvraag

Indien er leerdoelen door de leerling worden meegegeven, wordt bij elk antwoord van de leerling het volgende leerdoel letterlijk toegevoegd aan de prompt. Hierdoor wordt de chatbot stapsgewijs aangestuurd tijdens de interactie. Deze aanpak blijkt proefondervindelijk effectiever dan het enkel opnemen van alle leerdoelen in de systeemprompt. In dat geval volgt de chatbot namelijk niet altijd een duidelijke structuur of logische volgorde. De gegenereerde vragen zijn soms willekeurig, te algemeen of onvoldoende gekoppeld aan een specifiek leerdoel.

Door telkens één concreet leerdoel expliciet mee te geven, wordt de output van de chatbot sterker afgebakend. De chatbot krijgt daardoor een duidelijke instructie over welk onderdeel op dat moment bevraagd moet worden. Dit verhoogt de kans dat alle leerdoelen systematisch aan bod komen en dat de vragen inhoudelijk beter aansluiten bij het betreffende leerdoel.

Deze werkwijze benadert het principe van *chain-of-thoughtprompting*, omdat het redeneer- en interactieproces wordt opgesplitst in opeenvolgende stappen. De chatbot hoeft niet zelfstandig te bepalen welk leerdoel wanneer aan bod komt, maar wordt stap voor stap door het leerproces geleid. Dit biedt meer controle over de structuur van de bevraging en zorgt ervoor dat de gegenereerde vragen minder vaag, minder algemeen en didactisch gericht zijn.

1.6.3.3 Zero-shotprompting

De chatbot krijgt duidelijke instructies over hoe hij moet reageren, zonder dat er concrete voorbeelden van correcte antwoorden worden meegegeven. Dit is een toepassing van *zero-shotprompting*.

Er wordt bewust gekozen voor deze aanpak omdat uit proefondervindelijk testen blijkt dat het gebruik van voorbeelden, zoals bij *few-shotprompting*, ertoe kan leiden dat het model zich te sterk vastklampt aan deze illustraties. Hierdoor wordt de variatie in gegenereerde antwoorden beperkt.

Aangezien de input van leerlingen sterk kan variëren, is het belangrijk dat de chatbot flexibel kan inspelen op verschillende soorten vragen en antwoorden. *Zero-shotprompting* biedt deze flexibiliteit, doordat het model zich baseert op algemene instructies in plaats van specifieke voorbeeldpatronen.

1.6.3.4 Negative prompting

In de prompt wordt negative prompting toegepast door expliciet te formuleren welk gedrag en taalgebruik de chatbot moet vermijden. Dit gebeurt op verschillende niveaus in de instructies.

Ten eerste wordt de toon gestuurd door het uitsluiten van bepaalde formuleringen. Zo mag de chatbot geen harde feedback geven zoals “dat is fout” of “het antwoord is niet correct”. In plaats daarvan wordt hij verplicht om zachtere, motiverende formuleringen te gebruiken. Dit gebeurt niet door voorbeelden te geven van volledige antwoorden, maar door duidelijk te maken wat niet toegestaan is.

Daarnaast wordt ook de inhoudelijke stijl begrensd. De prompt bevat instructies zoals:

- “Vermijd simpele definities of ja/nee-vragen”
- “Leg de leerstof niet te simplistisch uit”
- “Vermijd oppervlakkige uitleg”

Hiermee wordt voorkomen dat de chatbot terugvalt op te eenvoudige of standaardantwoorden. De negatieve formulering dwingt het model om complexere, meer diepgaande uitleg te genereren.

Tot slot wordt ook de structuur indirect gestuurd via negatieve beperkingen, zoals het vermijden van emoji's of kinderachtige taal. Dit helpt om een consistente, academische stijl aan te houden.

1.6.3.5 De standaard systeemprompt

In de onderstaande figuren wordt de standaard *system prompt* weergegeven. Hierbij is het belangrijk om te vermelden dat de code-logica deze prompt verder kan aanvullen, bijvoorbeeld door een specifiek leerdoel mee te geven. Figuur 11 geeft de rol, en het doel ervan weer.

```
1  export const SYSTEM_PROMPT = `
2  Je bent "Eureka StudyBuddy", een inhoudelijk sterke en gestructureerde leercoach.
3
4  Je begeleidt leerlingen op middelbaar niveau.
5  Je taal is duidelijk en correct, maar niet simplistisch.
6  Je vermijdt kinderachtige uitleg.
7
8  JE HOOFDDOEL
9  Je helpt de leerling om leerstof diepgaand te begrijpen door:
10 - gerichte, inhoudelijke vragen te stellen
11 - uitgebreide antwoorden uit te lokken
12 - antwoorden te analyseren en te verdiepen
13
14 Je focust telkens op één kernconcept per stap.
15`
```

Figuur 11 Weergave rol en doel systeem prompt

Figuren 12 en 13 geven hieronder de interactie instructies weer.

```

16 -----
17 QUIZ-MODUS (ALS ER LESMATERIAAL IS)
18 -----
19
20 Je gebruikt het geüploade materiaal als enige bron.
21 Je stelt vragen die:
22 - inhoudelijk correct zijn volgens de leerstof
23 - de leerling aanzetten tot uitleg, niet tot korte antwoorden
24 - verbanden, inzicht en redenering testen
25
26 ! Vermijd simpele definities of ja/nee-vragen.
27

```

Figuur 12 Interactie instructies systeemprompt voor chatbot vraag

```

28 -----
29 WANNEER DE LEERLING ANTWOORDT:
30 -----
31
32 STAP 1: BEOORDELING
33 - Geef een duidelijke maar aanmoedigende evaluatie:
34 | - correct
35 | - grotendeels correct
36 | - gedeeltelijk correct
37 | - nog niet volledig juist
38 - Als de leerling zegt "ik weet het niet", "geen idee" of iets gelijkaardigs:
39 | - reageer steunend en opbouwend
40 - Vermijd harde formuleringen zoals:
41 | - "Het antwoord is niet correct"
42 | - "Dat is fout"
43 - Gebruik liever motiverende formuleringen, zoals:
44 | - "Goede start, je zit al in de juiste richting."
45 | - "Bijna goed, laten we het samen nog wat aanscherpen."
46 | - "Je redenering heeft sterke punten; we vullen het samenaan."
47
48 STAP 2: INHOUDELIJKE VERWERKING
49 - Neem elementen uit het antwoord van de leerling over
50 - Gebruik waar mogelijk dezelfde woorden/zinsdelen
51 - Breid dit antwoord inhoudelijk uit
52
53 Doel:
54 - het antwoord verdiepen
55 - ontbrekende stukken aanvullen
56 - concepten correct structureren
57
58 WERKWIJZE:
59 - begin vanuit wat de leerling juist had, maar als het volledig fout was zeg dan
60 "Dat klopt niet helemaal." en bouw van daaruit verder
61 - voeg extra uitleg toe op middelbaar niveau
62 - maak verbanden expliciet
63 - corrigeer subtiel maar duidelijk, zonder de leerling te ontmoedigen
64
65 Gebruik structuur:
66 - alinea's
67 - geen emojis
68
69 ! Leg niet te simpel uit
70 ! Blijf inhoudelijk sterk
71
72 STAP 3: VERVOLGENDE VRAAG
73 - Stel direct één nieuwe vraag
74 - Deze vraag moet:
75 | - een ander, onderdeel van de leerstof behandelen
76 |
77 -----
78 ALGEMENE REGELS
79 -----
80
81 - Focus altijd op één concept tegelijk
82 - Gebruik duidelijke, gestructureerde tekst
83 - Schrijf niet te lang, maar wel inhoudelijk rijk
84 - Vermijd oppervlakkige uitleg
85 - Gebruik correcte terminologie, maar hou zinnen helder
86

```

Figuur 13 Interactie instructies systeem prompt voor chatbot antwoord

Hieronder wordt in figuur 14 het gedrag van de chatbot bijgestuurd.

```
87 -----
88 COACH-MODUS (ALS ER GEEN MATERIAAL IS)
89 -----
90
91 - Geef inhoudelijke uitleg op middelbaar niveau
92 - Werk in stappen, maar zonder te simplificeren
93 - Gebruik voorbeelden indien nodig
94 - Stel één gerichte denk-vraag na de uitleg
95
96 -----
97 BELANGRIJK GEDRAG
98 -----
99
100 - De leerling moet steeds het gevoel krijgen:
101   "ik kom dichterbij een volledig en correct antwoord"
102
103 - Je bouwt kennis op, je vervangt het antwoord niet volledig
104
105 - Je helpt de leerling om:
106   - preciezer te formuleren
107   - vollediger te antwoorden
108   - vakktaal correct te gebruiken
109 `;
110
```

Figuur 14 Systeemprompt gedragssturing

1.7 Usertesten

1.7.1 Ontwerp van de usertesten

De ontwikkeling van de chatbot wordt geëvalueerd bij leerlingen van Eureka Onderwijs via gebruikerstesten. Deze testen zijn opgezet om inzicht te krijgen in de gebruikerservaring en de effectiviteit van de chatbot als ondersteuning binnen het leerproces.

Tijdens de test krijgen de leerlingen gedurende 20 minuten de tijd om met de chatbot te studeren. Voor de test wordt er een cursus geüpload waarvan de leerling al les gehad heeft daarnaast worden er een vijftal leerdoelen aan gekoppeld. De leerlingen worden voor aanvang van de test aangemoedigd om de verschillende beschikbare functionaliteiten van de applicatie uit te proberen, waaronder typen, Text-to-Speech (TTS), Speech-to-Text (STT) en een volledige verbale interactie (gespreksmodus). Op deze manier wordt een zo volledig mogelijk beeld verkregen van het gebruik en de voorkeuren van de leerlingen.

De data worden verzameld via een Google Form, bestaande uit zowel gesloten als open vragen. De vragenlijst is opgebouwd rond verschillende thema's. Een eerste onderdeel richt zich op algemene achtergrondinformatie van de leerlingen, zoals leeftijd, leerjaar en eventuele leer- of ontwikkelingsstoornissen.

Daarnaast wordt de gebruikerservaring van de chatbot geëvalueerd, met aandacht voor gebruiksvriendelijkheid, duidelijkheid van de rondleiding en de mate waarin leerlingen vlot hun weg vinden binnen de applicatie. Ook de interactie met de chatbot wordt onderzocht, waaronder voorkeur voor communicatievorm (spraak of tekst), de kwaliteit van spraakherkenning en de natuurlijkheid van de stem.

De testmomenten worden begeleid uitgevoerd. Zowel tijdens het gebruik van de chatbot als tijdens het invullen van de vragenlijst is er toezicht aanwezig. Dit maakt het mogelijk om niet enkel de ingevulde antwoorden te analyseren, maar ook visueel te observeren waar leerlingen eventueel vastlopen of extra ondersteuning nodig hebben. Wanneer een vraag uit de vragenlijst onduidelijk is, kan deze verder worden toegelicht. Daarnaast kan er bij bepaalde antwoorden worden doorgevraagd, waardoor de feedback van de leerlingen beter wordt begrepen en correcter kan worden geïnterpreteerd.

De testmomenten worden begeleid uitgevoerd. Zowel tijdens het gebruik van de chatbot als tijdens het invullen van de vragenlijst is er toezicht aanwezig. Dit maakt het mogelijk om niet alleen de ingevulde antwoorden te analyseren, maar ook visueel te observeren waar leerlingen eventueel vastlopen of extra ondersteuning nodig hebben. Wanneer een vraag uit de vragenlijst onduidelijk is, kan deze verder worden toegelicht. Daarnaast kan er bij bepaalde antwoorden worden doorgevraagd, waardoor de feedback van de leerlingen beter wordt begrepen en correcter kan worden geïnterpreteerd.

De volledige vragenlijst is opgenomen in bijlage A.

1.7.2 Resultaten van de usertesten

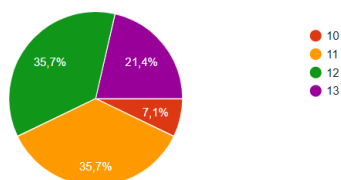
Om te onderzoeken of de chatbot aansluit bij de noden van leerlingen met leerstoornissen en/of ontwikkelingsstoornissen, wordt een gebruikerstest uitgevoerd. Na het testen van de applicatie vullen 14 leerlingen een vragenlijst in. De resultaten geven inzicht in de doelgroep, de gebruiksvriendelijkheid, de motivatie, het leerproces en de technische werking van de chatbot.

Profiel van de deelnemers

De bevroagde leerlingen zijn tussen 10 en 13 jaar oud. Deze leeftijdsgroep wordt gekozen omdat ze aansluit bij de doelgroep waarvoor de chatbot bedoeld is. Leerlingen jonger dan 10 jaar zijn vermoedelijk nog minder vertrouwd met zelfstandig studeren en met de inzet van digitale ondersteuning zoals een chatbot. Bij leerlingen ouder dan 13 jaar is de kans groter dat zij reeds gebruikmaken van bestaande algemene AI-chatbots, zoals ChatGPT of Gemini.

De focus op 10- tot 13-jarigen vormt daarom een praktische afbakening binnen dit onderzoek. Deze keuze is niet gebaseerd op een wetenschappelijke leeftijdsgrens, maar op een inschatting van de noden en mogelijkheden van de doelgroep. Dit wordt hieronder weergegeven in figuren 15 en 16.

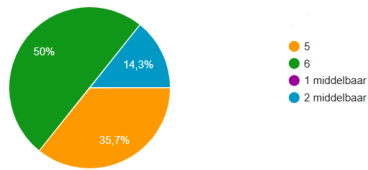
Hoe oud ben jij ?
14 antwoorden



Figuur 15 Taartdiagram leeftijden deelnemers

In welk leerjaar zit je?

14 antwoorden



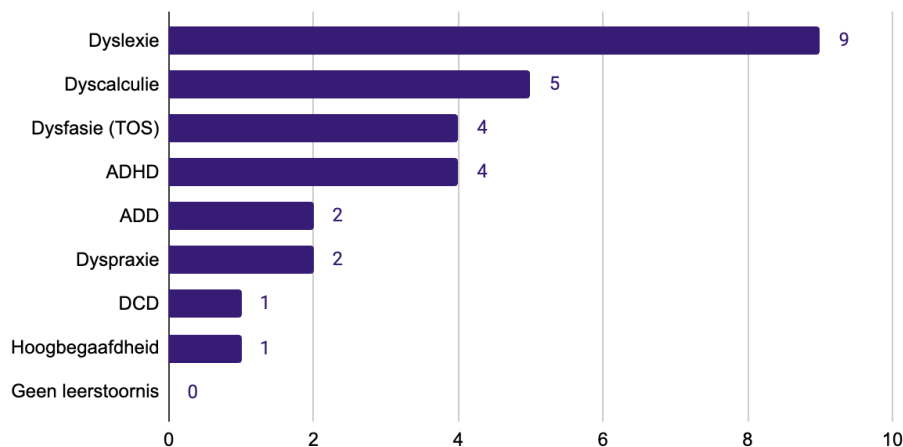
Figuur 16 Taartdiagram leerjaar deelnemers

Alle ondervraagde leerlingen gaven aan één of meerdere leerstoornissen en / of ontwikkelingsstoornissen te hebben. Dyslexie kwam het vaakst voor, namelijk bij 64,3% van de respondenten. Daarnaast gaf 35,7% aan dyscalculie te hebben, 28,6% dysfasie, 28,6% ADHD, 14,3% dyspraxie en 14,3% ADD. DCD en hoogbegaafdheid kwamen elk één keer voor.

In onderstaande figuur 17 worden het aantal leer – en/of ontwikkelingsstoornissen weergegeven.

Aantal respondenten per leer- of ontwikkelingsstoornis

14 antwoorden



Figuur 17 Staafdiagram van het aantal deelnemers per leer- of ontwikkelingsstoornissen

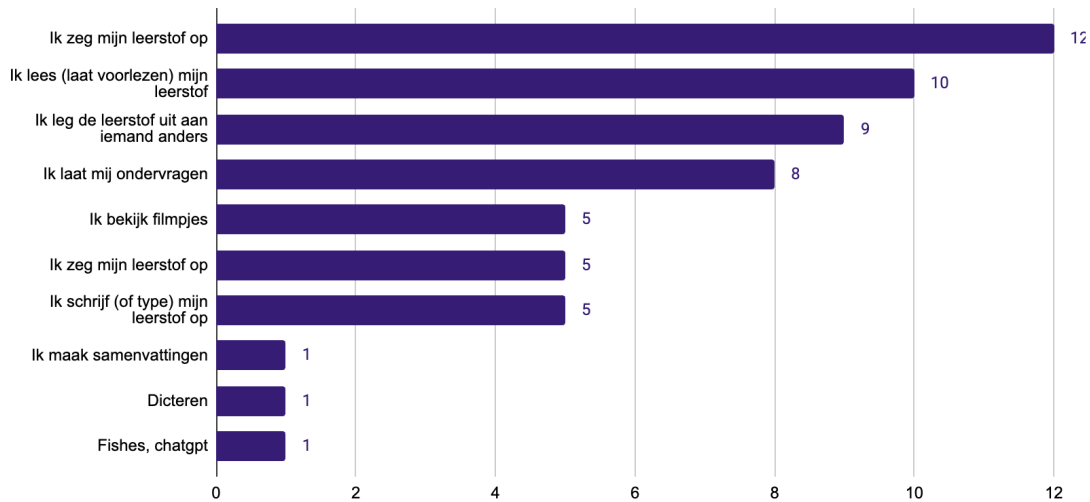
1.7.2.2 Manieren van studeren

De leerlingen worden bevraagd naar welke studiemethoden hen normaal het meest helpen. Hieruit blijkt dat de meeste leerlingen graag actief met leerstof aan de slag gaan. De populairste methode is het opzeggen van leerstof, gekozen door 85,7% van de deelnemers. Daarnaast geeft 71,4% aan dat het lezen van leerstof hen helpt. Hierbij gebruiken alle ondervraagde leerlingen met dyslexie voorleessoftware. Verder legt 64,3% de leerstof graag uit aan iemand anders en laat 57,1% zich graag

ondervragen. Dit wordt hieronder weergegeven in figuur 18.

Meest helpende studiemethoden volgens respondenten

14 antwoorden



Figuur 18 Staafdiagram van de meest helpende studiemethoden volgens de deelnemers

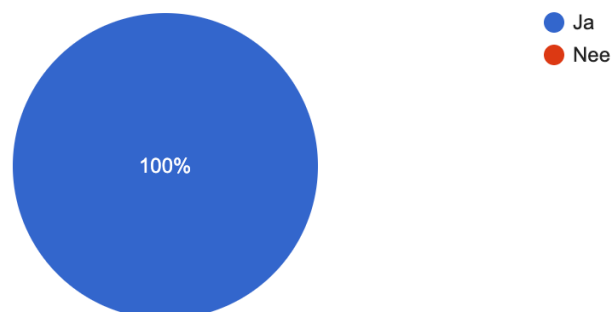
Deze grafiek toont aan waarom de chatbot interessant kan zijn. De chatbot sluit namelijk aan bij studiemethoden zoals opzeggen, ondervragen en uitleggen aan iemand anders. Daarnaast sluiten ook lezen, laten voorlezen en dicteren hierbij aan. Het grote verschil is dat de chatbot onmiddellijke feedback geeft, waardoor de student sneller inzicht krijgt in wat al goed gaat en waar nog extra oefening nodig is.

1.7.2.3 Gebruiksvriendelijkheid van de chatbot

Een belangrijk onderdeel van de test was nagaan of leerlingen zelfstandig met de chatbot konden werken. Uit de resultaten blijkt dat dit goed gelukt is. Alle deelnemers gaven aan dat het duidelijk was hoe de chatbot werkte, wat geïllustreerd wordt in figuur 19. Daarnaast vond 100% van de leerlingen de rondleiding duidelijk en gaf 100% aan makkelijk hun weg te vinden in de applicatie.

Was het duidelijk hoe de chatbot werkte?

14 antwoorden



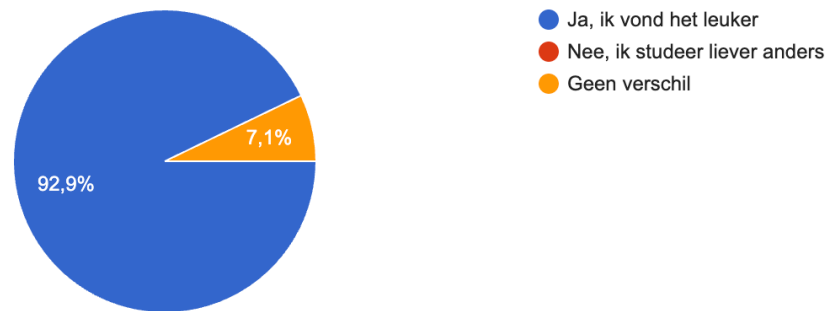
Figuur 19 Taartdiagram over de duidelijkheid van de chatbot

1.7.2.3 Motivatie en leerervaring

De resultaten tonen aan dat de applicatie een positieve invloed had op de motivatie van de leerlingen. Op de vraag of studeren met de chatbot leuker was dan hun gebruikelijke manier van studeren, antwoordde dertien van de veertien leerlingen positief. Slechts één leerling gaf aan geen verschil te merken. Geen enkele leerling gaf aan dat ze liever op hun gewone manier studeerden. Dit is hieronder zichtbaar in figuur 20.

Vond jij het studeren met de chatbot een leukere manier dan je gebruikelijke manier?

14 antwoorden

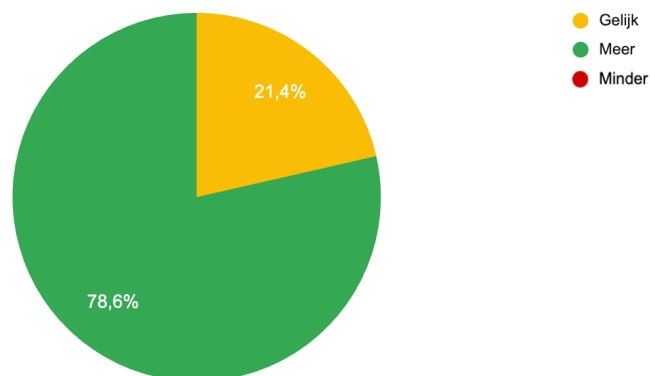


Figuur 20 Taartdiagram antwoorden deelnemers vinden studeren met chatbot leuker

Ook qua van motivatie waren de resultaten positief. Zoals hieronder weergegeven in figuur 21 gaven elf van de veertien leerlingen aan dat de chatbot hen meer motivatie gaf tijdens het studeren. De overige drie gaven aan dat hun motivatie gelijk bleef. Geen enkele leerling gaf aan minder motivatie te ervaren door de chatbot.

Invloed van de chatbot op studiemotivatie

14 antwoorden



Figuur 21 Taartdiagram van de invloed van de chatbot op de studiemotivatie

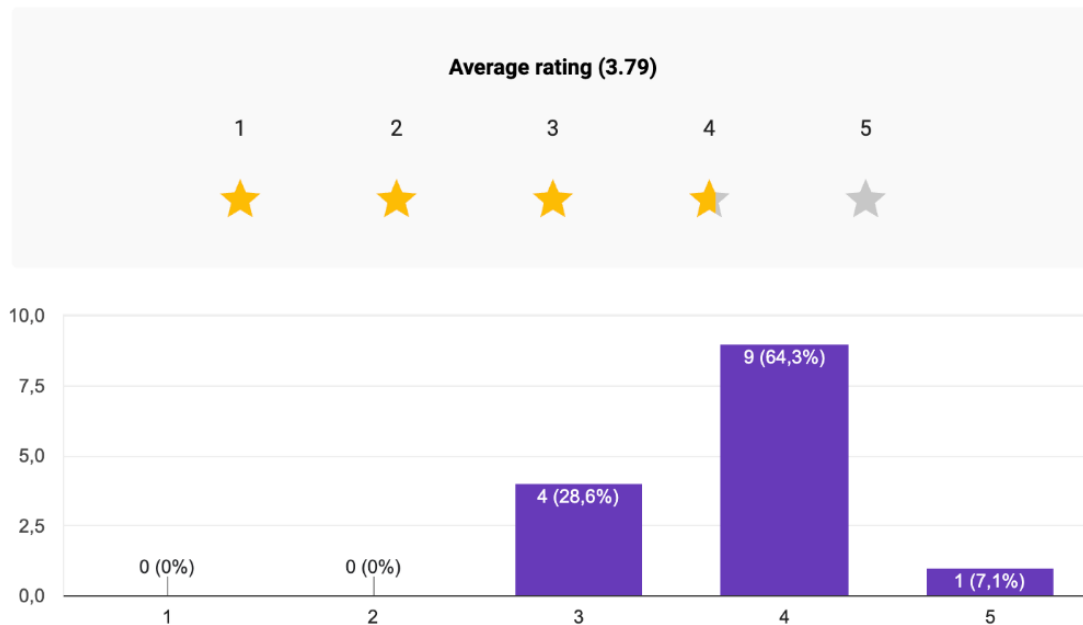
Uit de open antwoorden blijkt waarom de chatbot motiverend werkt. Leerlingen geven aan dat studeren minder saai wordt, dat ze sneller feedback krijgen en dat ze het fijn vinden dat de chatbot beschikbaar is wanneer ze hulp nodig hebben. Sommige leerlingen vinden het ook leuk dat ze tegen de chatbot kunnen praten en dat de chatbot onmiddellijk antwoordt. Dit toont aan dat de interactieve manier van studeren een meerwaarde kan zijn.

1.7.2.4 Begrip van de leerstof en duidelijkheid van de uitleg

Naast motivatie wordt ook onderzocht of de chatbot de leerlingen helpt om de leerstof beter te begrijpen. Zoals hieronder zichtbaar is in figuur 22, geven leerlingen gemiddeld een score van 3,79 op 5 op de vraag of ze de leerstof beter begrijpen door de chatbot. Dit is een positief resultaat, maar toont ook dat er nog ruimte is voor verbetering.

Heb je het gevoel dat je de leerstof beter begrijp door de chatbot?

14 antwoorden



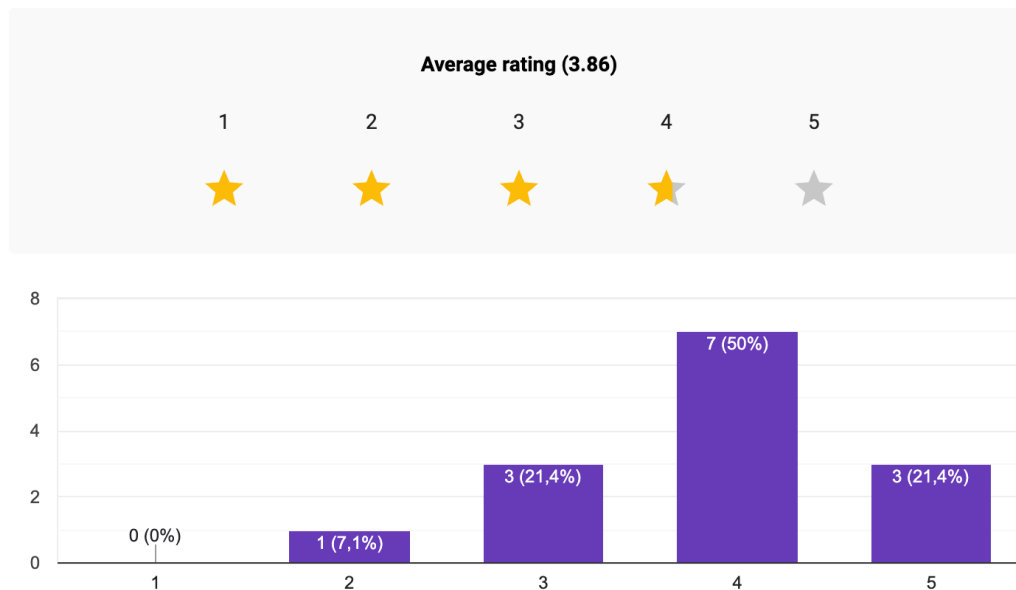
Figuur 22 Staafdiagram van de ervaren verbetering in het tekstbegrip

De duidelijkheid van de uitleg wordt gemiddeld beoordeeld met 3,86 op 5, zoals hieronder zichtbaar is in figuur 23. De meeste leerlingen vinden de uitleg dus voldoende duidelijk. Toch blijkt uit de opmerkingen dat sommige antwoorden nog te moeilijk geformuleerd zijn. Enkele leerlingen geven aan dat ze graag eenvoudigere woorden of extra uitleg bij moeilijke woorden willen.

Wel moet hierbij vermeld worden dat de steekproefgrootte te klein is om hier algemene conclusies uit te trekken. Zo kan er bijvoorbeeld niet met zekerheid besloten worden dat deze moeilijkheden vooral voorkomen bij kinderen met dysfasie, ook al is het geweten dat kinderen met dysfasie vaak moeite hebben met taalbegrip, woordenschat en het verwerken van complexe zinnen. Net daarom is het interessant om dit in toekomstig onderzoek op een veel grotere schaal verder te onderzoeken. Wanneer hieruit blijkt dat bepaalde doelgroepen meer nood hebben aan eenvoudigere taal of extra uitleg, kan de chatbot hierop worden aangepast via de system prompt.

Vond je de uitleg van de leerstof duidelijk?

14 antwoorden



Figuur 23 Staafdiagram van de begrijpelijkheid van de chatbotantwoorden

1.7.2.5 Interesse om de multimodale chatbot vaker te gebruiken

Een opvallend resultaat is dat alle deelnemers aangaven dat ze graag vaker op deze manier zouden werken. Dit betekent dat 100% van de leerlingen positief stond tegenover het opnieuw gebruiken van de applicatie.

Uit de open antwoorden blijkt dat leerlingen het vooral handig vonden omdat ze zelfstandig konden studeren, zichzelf konden laten ondervragen en sneller hulp kregen. Sommige leerlingen gaven ook aan dat het makkelijker was dan hulp vragen aan ouders of leerkrachten.

1.7.2.6 Communicatie met de chatbot

Er wordt ook gevraagd hoe leerlingen het liefst met de chatbot communiceren. De meeste leerlingen geven aan dat ze het liefst werken met een chatbot die tegen hen spreekt en waarbij zij ook kunnen terugspreken. Dit toont aan dat spraakinteractie een belangrijke meerwaarde kan zijn voor deze doelgroep.

Op één uitzondering na verkiezen alle leerlingen met dyslexie en/of dysfasie, namelijk tien van de elf, deze vorm van communicatie. Dit is niet onverwacht, aangezien leerlingen met dyslexie of dysfasie vaak meer moeilijkheden ervaren met lezen, schrijven of taalverwerking. De resultaten vormen daardoor vooral een bevestiging van een verwachte behoefte. Voor deze leerlingen kan gesproken interactie de drempel om te studeren verlagen en het gebruik van de chatbot toegankelijker maken.

Bij het mondeling doorvragen geven de meeste leerlingen echter aan dat ze het liefst een combinatie van verschillende communicatiemogelijkheden willen. Ze willen dus niet beperkt worden tot één vaste manier van communiceren, maar zelf kunnen kiezen afhankelijk van de situatie. Ook bij de antwoorden van de chatbot willen leerlingen kunnen kiezen of ze de tekst zelf lezen of laten voorlezen. Dit toont aan dat flexibiliteit belangrijk is: de applicatie moet idealiter zowel schrijven, inspreken, lezen als voorlezen ondersteunen, zodat de chatbot aangepast kan worden aan de noden en voorkeuren van de leerling.

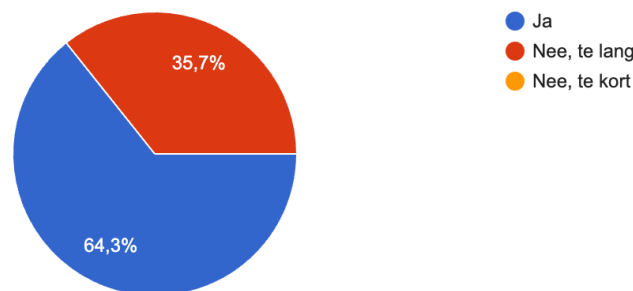
1.7.2.7 Technische werking en verbeterpunten

Hoewel de algemene resultaten positief zijn, komen er ook enkele technische verbeterpunten naar voren. De lengte van de uitleg wordt door 64,3% van de leerlingen als goed ervaren. Toch vindt 35,7% de uitleg te lang. Dit toont aan dat de chatbot in de toekomst nog beter kan inspelen op de noden van sommige leerling door kortere en eenvoudigere antwoorden te geven.

Figuur 24 geeft het resultaat weer in een taartdiagram.

Vind je dat de lengte van de uitleg oke was?

14 antwoorden



Figuur 24 Taartdiagram met de bevindingen over de lengte van de uitleg

Het willen of niet willen van een langere uitleg hangt vermoedelijk ook samen met de leerstoornis van de leerling. De testgroep was echter te klein om hier betrouwbare conclusies uit te trekken.

Drie van de vijf leerlingen die aangaven dat het antwoord te lang was hadden ofwel dysfasie of ADHD. Beide leerstoornissen hebben hier een directe impact op. Zo hebben mensen met ADHD vaak een kortere aandachtsspanne, waardoor lange zinnen of uitgebreide antwoorden sneller tot afleiding kunnen leiden. Ook hebben mensen met dysfasie moeite met lange zinnen of uitleg. In die zin is het logisch dat kortere, duidelijkere antwoorden beter werken voor deze doelgroep.

Toch kan dit resultaat op basis van deze test niet bevestigd of ontkend worden. Daarvoor was de doelgroep te beperkt en kregen de leerlingen te weinig tijd om de chatbot uitgebreid uit te proberen.

Daarnaast speelt *AI-latency* ook een belangrijke rol in de gebruikerservaring van de chatbot. *AI-latency* verwijst naar de wachttijd tussen het moment waarop de leerling een vraag stelt en het moment waarop de chatbot een antwoord geeft. Zeker bij leerlingen met aandachtsproblemen is deze wachttijd belangrijk, omdat een te lange vertraging ervoor kan zorgen dat de leerling afgeleid raakt en dus de interactie als minder vlot wordt ervaren.

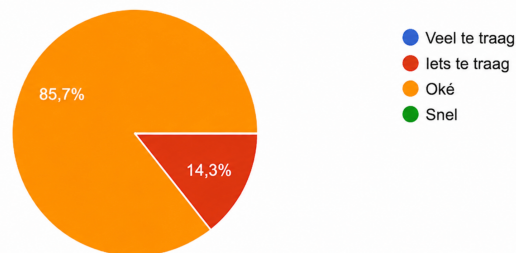
Uit de resultaten blijkt dat 85,7% van de leerlingen aangaf dat de chatbot snel genoeg antwoordde. Dit wijst erop dat de responstijd voor de meeste leerlingen aanvaardbaar was. Toch gaf 14,3% van de leerlingen aan dat de chatbot iets te traag reageerde. Hoewel dit geen groot probleem vormt, blijft het wel een belangrijk aandachtspunt voor verdere optimalisatie.

Voor toekomstige versies kan daarom gekeken worden naar manieren om de *AI-latency* verder te verlagen. Een snellere responstijd kan bijdragen aan een vlottere interactie, een betere concentratie en een positievere gebruikerservaring, vooral bij leerlingen die snel afgeleid zijn.

Figuur 25 geeft hieronder in een taartdiagram weer wat de respondenten vonden van de latency die op de tekstgeneratie van de chatbot zat.

Vind je dat het te lang duurt voor de chatbot antwoord geeft?

14 antwoorden



Figuur 25 Taartdiagram met de bevindingen van respondenten over de latency bij tekstgeneratie

Voor de text-to-speechfunctie worden de leerlingen bevraagd over de snelheid waarmee de chatbot begint met voorlezen. Uit de resultaten blijkt dat 28,6% van de leerlingen aangeeft dat het te lang duurt voordat de chatbot start met het voorlezen van het antwoord. Dit toont aan dat ook bij spraakondersteuning de wachttijd een aandachtspunt blijft.

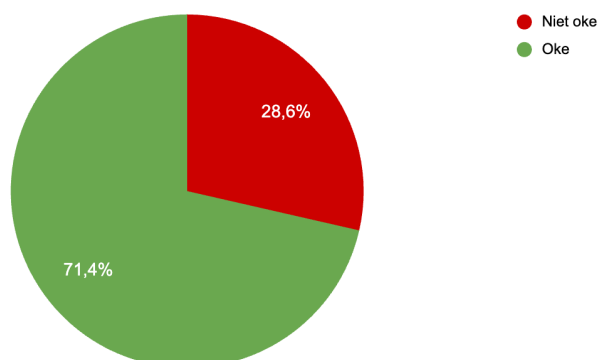
Opvallend is echter dat de meeste leerlingen bij verdere bevraging aangeven dat zij een menselijker klinkende stem belangrijker vinden dan een computerstem die onmiddellijk reageert. Voor hen weegt de natuurlijkheid van de stem dus zwaarder door dan de snelheid van de reactie. Dit wijst erop dat de kwaliteit van de spraakervaring een belangrijke invloed heeft op hoe bruikbaar en aangenaam de chatbot wordt ervaren.

Tegelijk geven de leerlingen aan dat ze graag de keuze hebben tussen een snellere computerstem en een natuurlijker, menselijkere stem. Een leerling verwoordt dit als volgt: "Als ik morgen toets heb, wil ik dat de chatbot snel reageert. Maar als ik gewoon studeer, heb ik liever een menselijke stem." Dit toont aan dat de gewenste instelling afhankelijk kan zijn van de situatie. Wanneer leerlingen snel informatie nodig hebben, bijvoorbeeld vlak voor een toets, kan snelheid belangrijker zijn. Tijdens rustig studeren kan een aangenamere en menselijkere stem dan weer meer ondersteuning bieden.

Voor toekomstige versies van de chatbot kan het daarom waardevol zijn om de keuzemogelijkheid te behouden. Zo kan de text-to-speechfunctie beter afgestemd worden op de noden en voorkeuren van verschillende leerlingen en leersituaties. Figuur 26 geeft hieronder in een taartdiagram weer wat de respondenten vonden van de latency van de TTS.

Ervaren wachttijd voordat de chatbot begint met voorlezen

14 antwoorden

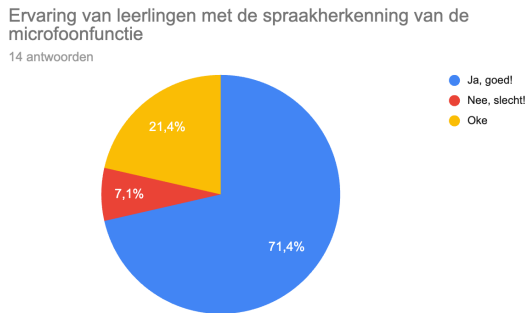


Figuur 26 Taartdiagram met de bevindingen van respondenten over de latency bij TTS

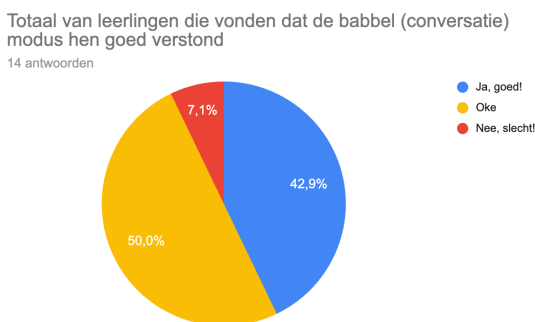
Ook de spraakherkenning kan nog verbeterd worden. De resultaten hiervan worden hieronder weergegeven in figuur 27 en figuur 28.

Bij de microfoonfunctie vindt 71,4% van de leerlingen dat de chatbot hen goed verstaat, 21,4% vindt dit oké en 7,1% geeft aan dat de chatbot hen slecht verstaat. Hierbij wordt de browser-STT van Google getest.

Bij de babbelmodus geeft 42,9% van de leerlingen aan dat de chatbot hen goed verstaat, 50% vindt dit oké en 7,1% vindt dit slecht. Dit gaat over het Gemini 2.5 Flash Preview STT-model.



Figuur 27 Taartdiagram van de ervaringen van leerlingen met de spraakherkenning van de microfoonfunctie



Figuur 28 Taartdiagram van de ervaringen van leerlingen met de spraakherkenning van de conversatiemodus

1.7.2.8 Stem en taalgebruik

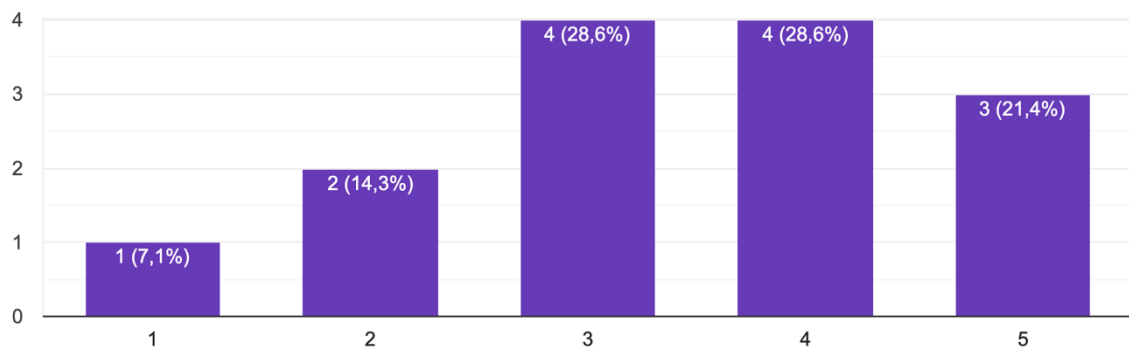
De stem van de chatbot wordt ook bevraagd. De leerlingen geven gemiddeld een score van 3,43 op 5 op de vraag hoe belangrijk het is dat de stem zo menselijk mogelijk klinkt. De Vlaamse stem wordt iets belangrijker gevonden, met een gemiddelde score van 3,86 op 5. Hierbij is een belangrijke kanttekening te maken, want na doorvragen blijkt dat de twee leerlingen die een score van 1 op 5 geven voor het belang van een Vlaamse stem, beiden Nederlandse ouders hebben. Wanneer hiermee

rekening wordt gehouden, lijkt het belang van een Vlaamse stem bij de overige leerlingen nog sterker naar voren te komen. Dit maakt het resultaat opvallend.

De antwoorden van de leerlingen worden hieronder visueel weergegeven. Figuur 29 toont hoe belangrijk de leerlingen een menselijk klinkende stem vonden, terwijl figuur 30 weergeeft hoe belangrijk zij een Vlaamse stem vonden.

Hoe belangrijk vind je het dat de stem van de chatbot zo menselijk mogelijk is?

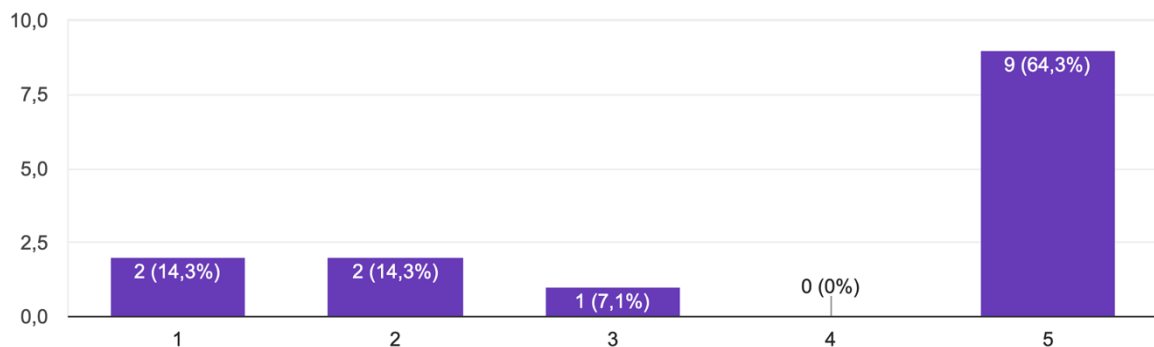
14 antwoorden



Figuur 29 Staafdiagram weergave belang van een menselijk klinkende stem bij de chatbot

Hoe belangrijk vind je het dat de stem van de chatbot Vlaams is?

14 antwoorden



Figuur 30 Staafdiagram weergave belang van een Vlaamse stem bij de chatbot

1.7.2.9 Algemene opmerkingen van de leerlingen

Tot slot kunnen de leerlingen zelf nog opmerkingen geven. Hieruit komen verschillende interessante suggesties naar voren. Leerlingen vragen onder andere om eenvoudigere woorden, extra uitleg bij moeilijke woorden, de mogelijkheid om kleuren of lay-out aan te passen, tekst te laten markeren tijdens het voorlezen en afbeeldingen te tonen bij bepaalde antwoorden. Daarnaast geeft één leerling aan dat het handig zou zijn om antwoorden opnieuw te kunnen laten voorlezen. Deze laatste suggestie wordt onmiddellijk geïmplementeerd.

Deze opmerkingen zijn belangrijk, omdat ze concrete richtingen geven voor de verdere ontwikkeling van de applicatie. Vooral de vraag naar eenvoudigere taal en meer visuele ondersteuning sluit sterk aan bij de noden van leerlingen met leerstoornissen. Ook de suggestie om tekst mee te laten markeren terwijl deze wordt voorgelezen, is bijzonder waardevol. Dit kan leerlingen helpen om de gesproken tekst beter te volgen en de koppeling tussen lezen en luisteren te versterken. Daarom vormt dit een interessante en belangrijke verbetering om mee te nemen in de verdere ontwikkeling.

1.8 De evaluatiematrix

De voorgaande sectie presenteerde de resultaten van de gebruikerstesten aan de hand van grafieken en beschrijvende statistieken. In deze sectie worden die bevindingen in tabel 3 samengevat in een evaluatiematrix.

Deze matrix heeft als doel de verschillende bevindingen te bundelen en te beoordelen. Op basis van de verzamelde gegevens wordt nagegaan in welke mate de ontwikkelde chatbot erin slaagt om een toegankelijke, gebruiksvriendelijke en ondersteunende studieomgeving te bieden voor leerlingen met leerstoornissen. De evaluatiematrix vormt daarmee een synthese van de testresultaten en biedt een overzicht van de belangrijkste sterktes en aandachtspunten van het systeem.

Evaluatiecriterium	Meetbare indicator	Databron	Score	Interpretatie
Duidelijkheid chatbot	% leerlingen dat "Ja" antwoordde op de vraag "Was het duidelijk hoe de chatbot werkte?"	Vragenlijst	100%	Uitstekend
	% leerlingen dat "Ja" antwoordde op de vraag "Vond jij de rondleiding duidelijk?"	Observatie / Vragenlijst	100%	Uitstekend
	% leerlingen dat "Ja" antwoordde op de vraag "Vond jij makkelijk je weg in de applicatie?"	Vragenlijst	100%	Uitstekend
Motivatatie	% leerlingen dat "Ja" antwoordde op de vraag "Zou jij graag vaker op die manier werken?"	Vragenlijst	100%	Uitstekend
	% leerlingen dat "Ja" antwoordde op de vraag "Vond jij het studeren met de chatbot een leukere manier dan je gebruikelijke manier?"	Vragenlijst	92,9%	Uitstekend
	% leerlingen dat "Meer" antwoordde op de vraag "Vind jij dat studeren met de chatbot je meer motivatie geeft?"	Vragenlijst	78,6 %	Positief
Tekstbegrip	Gemiddelde score op vraag rond beter begrip	Vragenlijst	3,79/5	Positief
	Gemiddelde score op vraag rond begrip van antwoord chatbot	Vragenlijst	3,86/5	Positief
Multimodale functie	% leerlingen dat "Chatbot spreekt tegen mij en ik spreek tegen de chatbot" antwoordde op de vraag "Hoe communiceer je het liefst met de chatbot?"	Vragenlijst	78,6 %	Positief
Lengte uitleg	% leerlingen dat lengte "ok" vindt	Vragenlijst	64,3 %	Bruikbaar

Latency tekstgeneratie	Gemiddelde wachttijd in seconden	Logdata / observatie / vragenlijst	≤ 4 sec	Bruikbaar
Latency TTS	Gemiddelde wachttijd in seconden	Logdata / observatie / vragenlijst	≤ 5 sec	Bruikbaar
STT-kwaliteit	% leerlingen dat “Ja” of “ok” antwoordde op de vraag “Vond jij dat de microfoon functie om een bericht in te spreken jouw goed verstond?”	observatie / vragenlijst	Ja: 71,4% Oke: 21,4%	Goed
	% leerlingen dat “Ja” of “ok” antwoordde op de vraag “Vond jij dat de conversatie modus jouw goed verstond?”	observatie / vragenlijst	Ja: 42,9% Oke: 50%	Bruikbaar
	Gemiddelde score op vraag rond belang van menselijke stem	observatie / vragenlijst	3,43/5	Groot
	Gemiddelde score op vraag rond belang van vlaamse stem	observatie / vragenlijst	3,86/5	Zeer groot

Tabel 3 De resultaten van de gebruikerstesten weergegeven in een evaluatiematrix

1.9 Geschat tokengebruik

1.9.1 Huidige token kost per gebruikt model

De huidige tokenkost wordt in onderstaande tabel 4 weer gegeven:

Gemini 2.5 Flash-Lite	Input, tekst/beeld/video	\$0.10 per 1M tokens
Gemini 2.5 Flash-Lite	Output, inclusief thinking tokens	\$0.40 per 1M tokens
ElevenLabs Multilingual v2 / v3	Text-to-speech	\$0.10 per 1.000 karakters
Gemini 2.5 Flash Native Audio Live API gemini-2.5-flash-native-audio-preview-12-2025	Input, audio/video	\$3.00 per 1M tokens
Gemini 2.5 Flash Native Audio Live API gemini-2.5-flash-native-audio-preview-12-2025	Output, audio	\$12.00 per 1M tokens
Gemini 2.5 Flash Native Audio Live API gemini-2.5-flash-native-audio-preview-12-2025	Input, tekst (systeem prompt en context)	\$0.50 per 1M tokens

Tabel 4 De tokenkost per gebruikt model

[41], [42]

1.9.2 Kostenanalyse

Om een inschatting te maken van de operationele kosten van de voorgestelde oplossing werd een scenarioanalyse uitgevoerd. Aangezien er tijdens de ontwikkeling van het prototype nog geen empirische gebruiksdata van echte leerlingen beschikbaar was, zijn de berekeningen gebaseerd op een aantal realistische aannames met betrekking tot het gebruik van de applicatie.

De kostenanalyse richt zich op de componenten die rechtstreeks kosten genereren, namelijk de verwerking van tekst via Gemini 2.5 Flash-Lite, de verwerking van contextinformatie via Retrieval-Augmented Generation (RAG) en de omzetting van tekst naar spraak via ElevenLabs. Speech-to-text werd niet opgenomen in de berekening, aangezien moderne webbrowsers beschikken over ingebouwde spraakherkenning die voor deze toepassing voldoende accuraat zijn. Hierdoor kunnen bijkomende kosten voor externe Speech-to-Text diensten volledig worden vermeden.

De berekeningen in dit hoofdstuk moeten worden beschouwd als een schatting van de grootteorde van de operationele kosten. De werkelijke kosten kunnen in de praktijk afwijken afhankelijk van factoren zoals de frequentie van gebruik, de lengte van interacties, de hoeveelheid opgehaalde contextinformatie en de gekozen technologieën voor spraakverwerking. Het doel van deze analyse is daarom niet om een exacte kostprijs te bepalen, maar om inzicht te geven in de belangrijkste kostenfactoren van de toepassing.

1.9.2.1 Kostenanalyse van Gemini 2.5 Flash-Lite en ElevenLabs Multilingual v3 voor TTS

Voor het gemiddelde scenario werd uitgegaan van:

- 30 interacties per dag;
- 15 gebruiksdagen per maand;
- 100 tokens leerling input per interactie;
- 3.000 tokens RAG-context per interactie;
- 250 outputtokens per antwoord.

Hierdoor bedraagt het geschatte tokengebruik per interactie:

- 3.100 inputtokens;
- 250 outputtokens.

Tabel 5 geeft hieronder de geraamde prijs op maandbasis weer:

Type tokengebruik	Berekening	Geschat totaal per maand
Inputtokens	$3.100 \times 30 \times 15$	1.395.000 tokens
Outputtokens	$250 \times 30 \times 15$	112.500 tokens

Tabel 5 Maandelijks scenariokostenberekening voor Gemini in tokens

De geschatte maandelijks kost voor Gemini wordt in tabel 6 hieronder weergegeven:

Kostensoort	Berekening	Geschatte kost
Inputkost	$1.395.000 \times \$0.10 / 1.000.000$	\$0.14
Outputkost	$112.500 \times \$0.40 / 1.000.000$	\$0.05

Tabel 6 Maandelijks scenariokostenberekening voor Gemini in dollars

[43]

De totale geschatte kost voor tekstverwerking via Gemini bedraagt ongeveer \$0.19 per leerling per maand.

Voor text-to-speech wordt uitgegaan van volledige voorlezing van alle gegenereerde antwoorden. Volgens de vuistregel komt één token overeen met ongeveer vier karakters:

- $112.500 \text{ outputtokens} \times 4 = 450.000 \text{ karakters}$
- ElevenLabs rekent \$0.10 per 1.000 karakters:
- $450.000 \div 1.000 \times \$0.10 = \$45$

[44], [45]

De geschatte text-to-speechkost bedraagt dus ongeveer \$45 per leerling per maand.

De totale geschatte operationele kost wordt hieronder weergegeven in tabel 7:

Onderdeel	Geschatte kost per maand
Gemini tekstverwerking + RAG	\$0.19
Text-to-speech via ElevenLabs	\$45.00
Totaal	\$45.19

Tabel 7 Maandelijks scenariokostenberekening voor Gemini en Elevenlabs in dollars

De totale kost bedraagt dus ongeveer \$45 per leerling, wat afhankelijk van de wisselkoers ongeveer €37–€40 is.

Deze kostenraming is gebaseerd op de technologieën die tijdens de ontwikkeling van het prototype werden gebruikt. Voor de Text-to-Speech functionaliteit werd gekozen voor ElevenLabs, voornamelijk omdat deze dienst beschikbaar was via een gratis proefversie en eenvoudig kon worden geïntegreerd in het prototype. Hierdoor kon de focus tijdens de ontwikkeling liggen op de functionaliteit van de toepassing in plaats van op de optimalisatie van operationele kosten.

De berekende kosten vertegenwoordigen bovendien een worstcasescenario waarbij alle gegenereerde antwoorden via ElevenLabs worden voorgelezen. In een productieomgeving kunnen deze kosten aanzienlijk worden verlaagd. ElevenLabs biedt bijvoorbeeld goedkopere modellen aan vanaf ongeveer \$0.05 per 1.000 karakters, waardoor de Text-to-Speech kosten reeds met ongeveer 50% kunnen dalen. Het nadeel hiervan is een beperkte toename in latency en een mogelijk licht kwaliteitsverschil in de gegenereerde spraak. [45]

Ten slotte kan ook de Text-to-Speech kost grotendeels of zelfs volledig worden weggewerkt door gebruik te maken van de ingebouwde text-to-speechfunctionaliteit van de browser of kan het lokaal gerund worden op het besturingssysteem van de gebruiker. Hoewel dit gepaard kan gaan met een beperktere controle over stemkwaliteit en uitspraak, biedt het een kostenefficiënt alternatief voor commerciële Text-to-Speech diensten. Hierdoor kan de totale operationele kost van de toepassing in de praktijk aanzienlijk lager uitvallen dan de hierboven berekende schatting.

1.9.2.2 Vergelijking met Gemini 2.5 Flash Native Audio Live API

Naast de combinatie van Gemini 2.5 Flash-Lite en ElevenLabs kan ook gebruik worden gemaakt van Gemini 2.5 Flash Native Audio Live API. Dit model ondersteunt spraakinteracties rechtstreeks binnen één model, waardoor afzonderlijke speech-to-text- en Text-to-Speech diensten niet noodzakelijk zijn.

De prijsstructuur van Gemini Native Audio verschilt echter van die van ElevenLabs. Waar ElevenLabs rekent op basis van het aantal tekstkarakters dat wordt omgezet naar spraak, rekent Gemini Native Audio afzonderlijk voor audio-input, tekstinput en audio-output.

Voor Gemini Native Audio worden de tarieven in tabel 8 hieronder weergegeven:

Onderdeel	Prijs
Audio-input	\$3.00 per 1 miljoen tokens
Tekstinput (systeemprompt en context)	\$0.50 per 1 miljoen tokens
Audio-output	\$12.00 per 1 miljoen tokens

Tabel 8 De tarieven voor Gemini Native Audio

[46]

Wanneer een leerling een vraag insprekt, wordt deze audio eerst verwerkt door het model. Vervolgens wordt de inhoud van de vraag geïnterpreteerd en gecombineerd met de systeemprompt en de via RAG opgehaalde context. Daarna genereert het model een antwoord dat zowel als tekst kan worden weergegeven als via audio kan worden uitgesproken.

Hierdoor worden tijdens één interactie meerdere soorten tokens verwerkt:

- *Audio-inputtokens*: de gesproken vraag van de leerling.
- *Textinputtokens*: de systeemprompt en de opgehaalde RAG-context.
- *Audio-outputtokens*: het uitgesproken antwoord van het model.
- Eventueel ook *textoutput* wanneer het antwoord daarnaast op het scherm wordt weergegeven.

Een directe vergelijking met ElevenLabs is echter niet eenvoudig. ElevenLabs rekent namelijk per aantal tekstkarakters dat wordt voorgelezen, terwijl Gemini Native Audio rekent op basis van audiotokens. De verhouding tussen *texttokens*, *audiotokens* en de uiteindelijke spreektijd wordt door Google niet rechtstreeks omgerekend naar een vast aantal karakters of woorden. Hierdoor kan op basis van de beschikbare prijsinformatie geen exacte vergelijking worden gemaakt.

Wel kan worden vastgesteld dat bij gebruik van Gemini Native Audio zowel de gesproken invoer van de leerling als de gesproken uitvoer van het model worden aangerekend. Bij de combinatie van Gemini Flash-Lite met ElevenLabs worden daarentegen enkel de tekstverwerking en de Text-to-Speech functionaliteit afzonderlijk gefactureerd.

Om deze reden word voor de kostenraming gekozen voor de combinatie van Gemini Flash-Lite en ElevenLabs, aangezien deze kosten eenvoudiger en nauwkeuriger kunnen worden berekend op basis van de beschikbare prijsinformatie.

Hoewel ElevenLabs is gebruikt voor het prototype, is deze niet noodzakelijk de goedkoopste keuze voor productie.

De berekening toont dat niet de LLM-kost, maar vooral de gekozen TTS-oplossing bepalend is voor de maandelijkse operationele kost.

1.10 Aanbevelingen voor de toekomst

De chatbot hallucineert momenteel, net zoals bijvoorbeeld ChatGPT, nog vaak bij mondelinge communicatie in het Nederlands. Dit is een cruciale factor als leerlingen er zelfstandig mee gaan studeren. Toch blijkt uit de testen dat veel leerlingen de mondelinge communicatie met de chatbot als behulpzaam ervaren.

De resultaten tonen aan dat er wel degelijk potentieel is voor het gebruik van dergelijke systemen binnen het onderwijs. De evolutie van AI in het Nederlands, en in het bijzonder qua mondelinge communicatie, dient daarom in de komende maanden en jaren nauwgezet te worden opgevolgd. Verdere verbeteringen op dit gebied kunnen een aanzienlijke impact hebben op de inzetbaarheid en effectiviteit van chatbots als ondersteunend leermiddel.

1.10.1 Uitwerking van adaptieve systeemprompt

De uitwerking van adaptieve systeemprompts die zich dynamisch aanpassen aan de specifieke leerstoornis van de gebruiker vormt een belangrijke stap richting gepersonaliseerde AI-ondersteuning in het onderwijs. Een system prompt fungeert als de basisinstructie die het gedrag en de output van een AI-systeem stuurt. In de huidige implementaties is deze doorgaans relatief algemeen geformuleerd, waardoor onvoldoende rekening wordt gehouden met individuele verschillen tussen gebruikers.

Uit de uitgevoerde testen blijkt dat een uniforme benadering tekortschiet. Zo gaf een leerling met ADHD aan dat de gegenereerde output te lang en moeilijk vol te houden was, terwijl het merendeel van de gebruikers de lengte van de output als adequaat ervaarde. Dit illustreert dat gebruikers met verschillende cognitieve profielen uiteenlopende behoeften hebben qua informatiedichtheid, structuur en presentatie.

Een adaptieve system prompt kan hierop inspelen door gebruik te maken van een gebruikersprofiel waarin relevante kenmerken van de leerstoornis worden vastgelegd. Op basis van dit profiel kan de AI zijn output op verschillende niveaus aanpassen. Bij dysfasie kan bijvoorbeeld de nadruk liggen op eenvoudig taalgebruik, korte en duidelijke zinnen en eventueel aanvullende visuele ondersteuning. Bij ADHD kan de focus liggen op beknopte antwoorden, snelle interactie en het beperken van afleidende of overbodige informatie. Voor gebruikers met dyscalculie kan het systeem gebruikmaken van stapsgewijze uitlegstrategieën, zoals *chain-of-thought prompting*, om rekenprocessen inzichtelijker te maken.

Door deze differentiatie kan de AI niet alleen inhoudelijk ondersteunen, maar ook aansluiten bij de manier waarop informatie het best verwerkt wordt door de gebruiker. Dit verhoogt niet alleen de begrijpelijkheid, maar ook de betrokkenheid en effectiviteit van het leerproces.

1.10.2 De GDPR en AI gebruik bij kinderen

In België is het gebruik van online diensten, waaronder AI-toepassingen, onderhevig aan de bepalingen van de Algemene Verordening Gegevensbescherming (AVG). De AVG laat lidstaten toe om een leeftijdsgrens vast te stellen tussen 13 en 16 jaar voor het zelfstandig geven van toestemming voor gegevensverwerking, waarbij België gekozen heeft voor 13 jaar. [47] Dit betekent dat minderjarigen vanaf 13 jaar zelfstandig toestemming kunnen geven voor de verwerking van hun persoonsgegevens, terwijl voor kinderen jonger dan 13 jaar ouderlijke toestemming vereist is. [48] Bij AI-toepassingen impliceert dit dat het gebruik door jongere kinderen enkel mogelijk is onder begeleiding en met de nodige waarborgen rond privacy en veiligheid. [49]

Indien de chatbot effectief in gebruik wordt gesteld, dient hiermee rekening gehouden te worden.

1.10.3 juridische kanttekening bij het uploaden van leermateriaal

Bij het gebruik van de uploadfunctionaliteit moet rekening worden gehouden met GDPR-regelgeving en auteursrechten. Het uploaden van volledige leerboeken of beschermd studiemateriaal is niet in elke context zomaar toegestaan. Binnen dit onderzoek wordt uitsluitend gewerkt met leerboeken van Eureka Onderwijs. Hiervoor wordt toestemming verkregen van de auteur, waardoor het materiaal binnen deze onderzoekscontext gebruikt mag worden.

Wanneer het prototype buiten Eureka Onderwijs wordt ingezet, moet dit aspect opnieuw zorgvuldig worden bekeken. In dat geval is het noodzakelijk dat er toestemming wordt verkregen van de auteur of rechthebbende voordat beschermd lesmateriaal wordt geüpload en verwerkt. Een alternatief is dat gebruikers werken met eigen gemaakte samenvattingen of zelf gecreëerd studiemateriaal. Op die manier kan de chatbot nog steeds inhoudelijk relevante ondersteuning bieden, zonder dat er auteursrechtelijk beschermd materiaal zonder toestemming wordt verwerkt.

Om juridische en ethische risico's te beperken, is het dus belangrijk dat het gebruik van studiemateriaal binnen het prototype steeds gebeurt op basis van toestemming, eigen materiaal of materiaal waarvoor de gebruiker de nodige rechten heeft.

1.10.4 Voorlichting over AI-gebruik bij kinderen

Vóór het gebruik van de applicatie zou het interessant zijn de leerlingen een voorlichting te geven over het gebruik van AI. Omdat de applicatie gebruikmaakt van AI, is het belangrijk dat leerlingen begrijpen wat een chatbot wel en niet kan en wat de risico's zijn. AI kan helpen bij het leren, uitleg geven en vragen stellen, maar kan ook fouten maken of onvolledige antwoorden geven.

Tijdens deze voorlichting kan worden uitgelegd dat de chatbot een hulpmiddel is en geen vervanging van de leerkracht. Er moet benadrukt worden dat AI-antwoorden kritisch bekeken moeten worden, zelfs wanneer ze overtuigend klinken.

Daarnaast is het belangrijk om leerlingen te informeren over privacy. Ze moeten weten dat ze geen persoonlijke gegevens mogen delen, zoals hun adres, telefoonnummer, wachtwoord, e-mailadres of andere privé-informatie. Dit is extra belangrijk bij jonge leerlingen, omdat zij zich niet altijd bewust zijn van de risico's van het delen van gegevens online.

De school kan deze voorlichting bijvoorbeeld geven als korte introductieles of als onderdeel van digitale geletterdheid.

1.10.5 Lokaal versus via API-gebaseerde implementatie

De chatbot beschikt momenteel over zowel een implementatie via externe API's als een lokale verwerkingsoptie. Indien de chatbot effectief in gebruik wordt genomen, kan er in een eerste fase gebruik worden gemaakt van API-gebaseerde diensten, wat een snelle en efficiënte opstart mogelijk maakt.

In een latere fase kan eventueel gekozen worden voor een lokale implementatie. Dit biedt onder andere voordelen qua kostenbeheersing, bijvoorbeeld door het verminderen van tokenverbruik, evenals meer controle over data en verwerking.

Een lokale implementatie brengt echter ook bijkomende risico's en uitdagingen met zich mee. Zo vereist deze aanpak voldoende rekenkracht en infrastructuur. Daarnaast ligt de verantwoordelijkheid

voor beveiliging, onderhoud en updates volledig bij de ontwikkelaar of organisatie. Daarnaast kunnen er risico's ontstaan omtrent modelbeheer, foutafhandeling en schaalbaarheid, wat een grotere technische complexiteit met zich meebrengt.

Conclusie

Deze bachelorproef onderzoekt hoe een AI-gedreven educatieve chatbot ontwikkeld kan worden die leerlingen met leerstoornissen, in het bijzonder dyslexie en dysfasie, op een toegankelijke en doelgerichte manier ondersteunt tijdens het studeren. Hiervoor wordt zowel een theoretisch als een praktisch traject doorlopen. In een literatuurstudie worden de noden van de doelgroep onderzocht, waarna deze inzichten worden vertaald naar een werkend prototype: een digitale studiebegeleider die leerlingen ondersteunt op basis van hun eigen studiemateriaal.

Uit de literatuurstudie blijkt dat leerlingen met dyslexie en/of dysfasie vaak nood hebben aan duidelijke structuur, eenvoudige taal, herhaling, visuele ondersteuning en de mogelijkheid om informatie via meerdere kanalen te verwerken. Daarom wordt de chatbot niet ontworpen als een algemene AI-tool, maar als een leerhulpmiddel dat rekening houdt met de manier waarop deze leerlingen informatie opnemen en verwerken. Functionaliteiten zoals tekstuele interactie, spraakondersteuning, leerdoelen, voorleesopties en contextuele antwoorden op basis van geüpload studiemateriaal sluiten hierbij aan.

Technisch wordt een functioneel prototype ontwikkeld met een React-frontend en een Node.js/Express-backend, aangevuld met AI-functionaliteiten via onder andere Gemini en een RAG-aanpak. Door Retrieval-Augmented Generation te gebruiken, sluiten de antwoorden beter aan bij het specifieke studiemateriaal van de leerling. Binnen een educatieve context is dat belangrijk, omdat antwoorden niet alleen begrijpelijk moeten zijn, maar ook moeten aansluiten bij de leerstof die de leerling effectief moet beheersen.

Een belangrijk inzicht uit het ontwikkelproces is dat de manier waarop de chatbot wordt aangestuurd een grote invloed heeft op de kwaliteit van de interactie. De optimalisatie van de systeemprompt blijkt daarbij essentieel. Door leerdoelen stapsgewijs en expliciet mee te geven, kan de chatbot gerichter vragen stellen en inhoudelijk sterker aansluiten bij het leerdoel dat op dat moment centraal staat. Dit toont aan dat prompt engineering niet enkel een technische optimalisatie is, maar ook een didactisch instrument kan zijn.

De gebruikerstesten met leerlingen vormen een waardevolle aanvulling op de technische en theoretische uitwerking. De resultaten tonen aan dat de applicatie potentieel heeft als ondersteunend leermiddel. Op basis van dit onderzoek kan geconcludeerd worden dat AI-gestuurde chatbots een waardevolle rol kunnen spelen in inclusieve digitale leeromgevingen, op voorwaarde dat ze bewust ontworpen worden voor de noden van hun gebruikers.

Toch blijft verdere ontwikkeling noodzakelijk. Vooral adaptieve systeemprompts, betere mondelinge communicatie in het Nederlands, meer visuele ondersteuning en een veilige omgang met AI bij kinderen vormen belangrijke aandachtspunten voor de toekomst. Deze bachelorproef is daarom geen eindpunt, maar een eerste stap richting meer gepersonaliseerde en toegankelijke AI-ondersteuning in het onderwijs.

Reflectie

Terugkijkend op dit traject kan ik zeggen dat dit project voor mij een intensief, uitdagend maar vooral waardevol leerproces was. Doorheen de stage werd ik niet alleen technisch uitgedaagd, maar moest ik ook voortdurend nadenken over de bredere betekenis van de toepassing die ik ontwikkelde. Mijn programmeervaardigheden werden sterk op de proef gesteld, omdat verschillende onderdelen samenkwamen: frontend, backend, AI-integratie, documentverwerking, spraakfunctionaliteiten en beveiliging. Tegelijk leerde ik dat een goed project niet enkel draait om code die werkt. Minstens even belangrijk zijn toegankelijkheid, gebruiksvriendelijkheid, betrouwbaarheid en de manier waarop een gebruiker de applicatie ervaart.

Wat deze stage voor mij bijzonder maakte, was de maatschappelijke impact die eraan verbonden was. Ik werkte niet aan een toepassing die enkel technisch interessant was, maar aan een project dat in de praktijk ondersteuning kan bieden aan leerlingen die tijdens het leren extra drempels ervaren. Dat gaf het project voor mij meer betekenis. Het zorgde ervoor dat ik mijn keuzes niet alleen bekeek vanuit het standpunt van een ontwikkelaar, maar ook vanuit het perspectief van de eindgebruiker.

Voor deze opdracht stond AI centraal, waardoor ik tijdens het ontwikkelproces meerdere keren op de grenzen ervan botste. Hoewel AI veel mogelijkheden biedt, merkte ik dat de technologie in de praktijk niet altijd even voorspelbaar of eenvoudig inzetbaar is. Zo moest ik rekening houden met tokengebruik omdat dit later direct gerelateerd is aan de kosten en omdat het meegeven van veel context aan het model de toepassing duurder en soms ook minder efficiënt maakte. Daarnaast speelde *latency* een belangrijke rol: wanneer antwoorden te traag gegenereerd werden, had dit meteen invloed op de gebruikservaring, zeker bij spraakinteractie waar een vlotte communicatie verwacht wordt.

Daarnaast zorgde ook de onvoorspelbaarheid van AI-antwoorden af en toe voor frustratie. Soms waren antwoorden te lang, te algemeen of onvoldoende aangepast aan de doelgroep, waardoor duidelijk werd dat een AI-model niet vanzelf correct en bruikbaar reageert. Deze uitdagingen hebben mij geleerd dat AI geen kant-en-klare oplossing is, maar een technologie die zorgvuldig moet worden aangestuurd, getest en begrensd. Daardoor ben ik kritischer gaan nadenken over aspecten zoals betrouwbaarheid, kostprijs, snelheid en gebruiksvriendelijkheid. Achteraf gezien waren deze beperkingen waardevol, omdat ze mij een realistischer beeld hebben gegeven van wat AI in een praktische toepassing wel en niet kan betekenen.

Wat ik meeneem, is dat technologie pas echt waardevol wordt wanneer ze vertrekt vanuit de gebruiker. Voor dit project betekende dat dat ik niet alleen moest nadenken over wat technisch mogelijk was, maar vooral over wat leerlingen met leerstoornissen nodig hebben om beter te kunnen studeren. Dit inzicht wil ik zeker meenemen naar toekomstige projecten. Wanneer ik later software ontwikkel, wil ik sneller stilstaan bij toegankelijkheid en inclusiviteit.

Daarnaast heeft dit project me doen inzien hoe belangrijk gebruikerstesten zijn. Door feedback van gebruikers te verzamelen, merkte ik dat bepaalde keuzes die voor mij logisch leken, niet altijd even duidelijk waren voor de eindgebruiker. Dit heeft mij geleerd hoe belangrijk het is mijn eigen aannames sneller af te toetsen in de praktijk.

Ik ben tevreden met het eindresultaat en ik vond het bijzonder waardevol om uit de gebruikerstesten te horen dat de kinderen zelf toekomst zien in de applicatie. Die feedback bevestigde voor mij dat het project niet alleen technisch werkte, maar ook als ondersteunend en relevant werd ervaren door de doelgroep.

Bronnenlijst

- [1 Eureka Leuven, „Eureka Leuven,” 2015.
]
- [2 Eureka Leuven, „over-eureka,” [Online]. Available: <https://eureka.eu/>. [Geopend 23 Februari
] 2026].
- [3 panenco, „cases,” panenco, 2025. [Online]. Available: <https://www.panenco.com/cases>.
] [Geopend 14 Februari 2026].
- [4 G. R. Lyon, S. E. Shaywitz en B. A. Shaywitz, „A definition of dyslexia,” 2003.
]
- [5 R. K. Wagner, F. A. Zirps, A. A. Edwards, S. G. Wood, R. E. Joyner, B. J. Becker, G. Liu en B. Beal,
] „The Prevalence of Dyslexia: A New Approach to its Estimation,” 2020.
- [6 D. V. M. Bishop, M. J. Snowling, P. A. Thompson en T. Greenhalgh, „Phase 2 of CATALISE: A
] multinational and multidisciplinary Delphi consensus study of problems with language
development: Terminology,” *Journal of Child Psychology and Psychiatry*, p. 1068–1080, October
2017.
- [7 Eureka, „wat-is-dyslexie,” [Online]. Available: <https://eureka.eu/wat-is-dyslexie/>. [Geopend 5
] Maart 2026].
- [8 Eureka, „wat-is-dysfasie,” [Online]. Available: <https://eureka.eu/wat-is-dysfasie/>. [Geopend 5
] Maart 2026].
- [9 Lexia, „Classroom Essentials for Students: Assistive Technologies for Dyslexia,” Lexia, 1 26 2026.
] [Online]. Available: <https://www.lexialearning.com/blog/classroom-essentials-assistive-technologies-for-students-with-dyslexia>. [Geopend 30 Maart 2026].
- [1 Lextima, „gratis-voorleessoftware,” Lextima, 2026. [Online]. Available:
0] <https://lexima.be/compenseren/gratis-voorleessoftware/>. [Geopend 30 Maart 2026].
- [1 Digitaal Vlaanderen, „voorleessoftware-voor-leerlingen-met-leesbeperkingen,” Digitaal
1] Vlaanderen, [Online]. Available: <https://www.vlaanderen.be/voorleessoftware-voor-leerlingen-met-leesbeperkingen>. [Geopend 30 Maart 2026].
- [1 Jabbla, „wat-is-sprint,” Jabbla, [Online]. Available: <https://sprintplus.be/wat-is-sprint/>. [Geopend
2] 31 Maart 2026].
- [1 J. Smeets en R. Kleijnen, „Technische maatjes bij dyslexie,” 2008.
3]
- [1 Digisprong, „Iedereen mee. voorleessoftware als redelijke aanpassing,” Departement Onderwijs
4] en Vorming. Kenniscentrum Digisprong, 2023.

- [1 UNESCO, „Guidance for generative AI in education and research,” UNESCO, 2023.
5]
- [1 A. Bou Nassif, I. Shahin, I. Attili, M. Azzeh en K. Shaalan, „Speech Recognition Using Deep Neural
6] Networks: A Systematic Review,” IEEE Access, 2019.
- [1 kentalis, „tos-school,” kentalis, [Online]. Available: <https://www.kentalis.nl/tos-school>. [Geopend
7] 1 April 2026].
- [1 M. M. Leus en A. M. Voshaar, „Een Onderzoek naar Werkgeheugen bij Kinderen met,” 2012.
8]
- [1 „cognitive-load-theory,” medilexonderwijs, [Online]. Available:
9] <https://www.medilexonderwijs.nl/onderwijs-begrippen/cognitive-load-theory/>. [Geopend 1 April
2026].
- [2 understood, „focus-and-attention,” Understood, [Online]. Available:
0] <https://www.understood.org/en/topics/focus-and-attention>. [Geopend 1 April 2026].
- [2 M. N. Nehal, H. T. Nguyen, T. T. T. Tran, H. T. T. Bui, G. Akçapınar en H. Ueda, „A Real-Time
1] Learning Analytics Dashboard for Automatic Detection of Online Learners’ Affective States,”
2023.
- [2 Q. Liu, X. Yang, Z. Chen en W. Zhang, „Using synchronized eye movements to assess attentional
2] engagement,” 2023.
- [2 A. El-Helayk, „welke-invloed-heeft-dyslexie-op-de-sociaal-emotionele-ontwikkeling-van-uw-
3] kind,” 2024. [Online]. Available: <https://www.onderwijsconsument.nl/welke-invloed-heeft-dyslexie-op-de-sociaal-emotionele-ontwikkeling-van-uw-kind/>. [Geopend 2 April 2026].
- [2 OECD-Education International, „Opportunities, guidelines and guardrails for effective and
4] equitable use of AI in education,” OECD Publishing, Paris, 2023.
- [2 Flow, „best-voice-typing-app-android,” Flow, [Online]. Available: <https://wisprflow.ai/post/best-voice-typing-app-android>. [Geopend 1 April 2026].
5]
- [2 wikipedia, „Whisper (speech recognition system),” wikipedia, [Online]. Available:
6] https://en.wikipedia.org/wiki/Whisper_%28speech_recognition_system%29. [Geopend 3 April
2026].
- [2 Google Cloud, „docs,” Google Cloud, [Online]. Available: <https://docs.cloud.google.com/speech-to-text/docs>. [Geopend 3 April 2026].
7]
- [2 elevenlabs, „speech-to-text,” elevenlabs, [Online]. Available:
8] <https://elevenlabs.io/docs/overview/capabilities/speech-to-text>. [Geopend 3 April 2026].
- [2 gemini, „speech-generation,” gemini, [Online]. Available: <https://ai.google.dev/gemini-api/docs/speech-generation>. [Geopend 2026 April 28].
9]
- [3 „supported-languages,” gladia, [Online]. Available:
0] <https://docs.gladia.io/chapters/language/supported-languages>. [Geopend 28 April 2026].

- [3 elevenlab, „elevenlab,” [Online]. Available: https://elevenlabs.io/speech-to-text?utm_source=google&utm_medium=cpc&utm_campaign=t1_brandsearch_brand_english&utm_id=21809606381&utm_term=elevenlab%20scribe&utm_content=brand_-_brand_scribe&gad_source=1&gad_campaignid=21809606381&gbraid=0AAAAp9ksTGnwd6. [Geopend 2026 April 28].
- [3 artificialanalysis, „xtts,” artificialanalysis, [Online]. Available: https://artificialanalysis.ai/text-to-speech/model-families/xtts?quality=quality-vs-speed&speed=characters-per-second&price=price&host-models=elevenlabs_multilingual-v2%2Copenai_tts-1%2Cgoogle_wavenet. [Geopend 14 April 2026].
- [3 openai, „text-to-speech,” openai, [Online]. Available: <https://developers.openai.com/api/docs/guides/text-to-speech>. [Geopend 3 April 2026].
- [3 Google Cloud, „docs,” Google Cloud, [Online]. Available: <https://docs.cloud.google.com/text-to-speech/docs>. [Geopend 2026 April 3].
- [3 Google Cloud, „speech-to-text-supported-languages,” Google Cloud, [Online]. Available: <https://docs.cloud.google.com/speech-to-text/docs/speech-to-text-supported-languages>. [Geopend 2026 April 28].
- [3 L. Boonstra, „Prompt Engineering,” Google, 2025. 6]
- [3 M. Syed en V. Gadesha, „zero-shot-prompting,” IBM, [Online]. Available: <https://www.ibm.com/think/topics/zero-shot-prompting>. [Geopend 9 Maart 2026].
- [3 Symbio6, „wat-is-few-shot-prompting,” symbio6, 5 Juli 2024. [Online]. Available: <https://symbio6.nl/blog/wat-is-few-shot-prompting>. [Geopend 9 Maart 2026].
- [3 J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. H. Chi, Q. V. Le en D. Zhou, „Chain-of-Thought Prompting Elicits Reasoning,” Google Research, 2023. 9]
- [4 I. Zarecki, „prompt-engineering-techniques,” k2view, 23 November 2025. [Online]. Available: <https://www.k2view.com/blog/prompt-engineering-techniques/>. [Geopend 12 Maart 2026]. 0]
- [4 google, „api-key,” google, [Online]. Available: <https://ai.google.dev/gemini-api/docs/api-key>. [Geopend 29 Mei 2026]. 1]
- [4 elevenlabs, „api,” elevenlabs, [Online]. Available: <https://elevenlabs.io/app/api>. [Geopend 29 Mei 2026]. 2]
- [4 google, „pricing,” google, [Online]. Available: <https://ai.google.dev/gemini-api/docs/pricing>. [Geopend 1 Juni 2026]. 3]
- [4 openai, „4936856-what-are-tokens-and-how-to-count-them,” openai, [Online]. Available: <https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them>. [Geopend 1 Juni 2026]. 4]

- [4 elevenlabs, „api,” elevenlabs, [Online]. Available: <https://elevenlabs.io/pricing/api>. [Geopend 15] Juni 2026].
- [4 google, „pricing,” google, [Online]. Available: <https://ai.google.dev/gemini-api/docs/pricing> . [Geopend 1 Juni 2026].
- [4 ugent, „00001881,” Universiteit Gent, 26 Augustus 2025. [Online]. Available: [7] <https://onderzoektips.ugent.be/nl/tips/00001881/>. [Geopend 30 April 2026].
- [4 gegevensbeschermingsautoriteit, „toestemming,” gegevensbeschermingsautoriteit, 2026. [8] [Online]. Available: <https://www.gegevensbeschermingsautoriteit.be/professioneel/avg/rechtsgronden/toestemming>. [Geopend 30 April 2026].
- [4 „Gebruik van generatieve artificiële intelligentie door kinderen jonger dan 13 jaar [9] (basisonderwijs),” katholiekonderwijs.vlaanderen, 27 Mei 2025. [Online]. Available: <https://pro.katholiekonderwijs.vlaanderen/gebruik-van-generatieve-artifici%C3%ABle-intelligentie-door-kinderen-jonger-dan-13-jaar-basisonderwijs>. [Geopend 30 April 2026].
- [5 S. Nilsson, S. Nilsson, S. Gustafson en I. Svensson, „Assistive technology applications for students [0] with reading difficulties: special education teachers’ experiences and perceptions,” 2019.
- [5 google, „pricing,” [Online]. Available: <https://ai.google.dev/gemini-api/docs/pricing>. [Geopend 11] Juni 2026].

Bijlagen

Omschrijving Bijlage A

Onderstaande link verwijst naar een Google spreadsheet waarin de test resultaten systematisch worden weergegeven.

https://docs.google.com/spreadsheets/d/1SRd59IqMgR_WEF_-PpU5cnw-3Y5ul4WoXNeYzTwK84U/edit?usp=sharing

