

---

# **EXTENDED ABSTRACT**

## **DI.ON.NE**

---

### **Mbo stage**

### **Zuyd Hogeschool**

Tygo Offermanns

Stakeholder: Marcel Schmitz

Begeleiders: Shannen Dolls, Jeroen Kornips

2024/2025  
11-07-2025

#### **Abstract**

De opkomst van kunstmatige intelligentie en spraakgestuurde systemen biedt nieuwe mogelijkheden voor gepersonaliseerde leeromgevingen. Naar aanleiding hiervan is een digitale avatar ontwikkeld die studenten ondersteunt bij het stellen en beantwoorden van vragen. De doelstelling was het ontwerpen en implementeren van een interactieve AI-assistent die toegankelijk is via zowel tekstuele als gesproken input. Hiervoor is een Unity-applicatie gebouwd waarin verschillende technologieën samenkommen: spraakherkenning (Speech-to-Text), natuurlijke taalverwerking met behulp van een lokaal Large Language Model (LLM via Ollama), en spraaksynthese (Text-to-Speech). De methode bestond uit het integreren van deze componenten in een gebruiksvriendelijke interface, gecombineerd met een modulaire backend in Flask. De resultaten tonen aan dat het systeem gebruikersinvoer effectief verwerkt en relevante antwoorden genereert in realtime. Gebruikers kunnen moeiteloos wisselen tussen spraak- en tekstmodus, wat bijdraagt aan de toegankelijkheid en gebruikservaring. Concluderend kan worden gesteld dat de toepassing van AI in dit systeem bijdraagt aan een intuïtieve en laagdrempelige leerondersteuning, wat relevant is voor toekomstig onderwijs en verdere ontwikkeling van digitale assistenten.

# 1 Inleiding

## 1.1 Aanleiding

Binnen dit project vervult de stakeholder de rol van product owner voor het geven van de requirements. Later heeft de stakeholder de rol van product owner overgedragen aan de begeleiders van het project.

Het DI-Lab is een onderdeel van het lectoraat Data Intelligence, een innovatieve werkplek die bestaat uit docent-onderzoekers van Hogeschool Zuyd[8]. Het DI-Lab bestaat uit veel specialisaties, onder andere AI (Artificial Intelligence). Deze specialisatie heeft veel te maken met het DI.ON.NE project, aangezien de avatar gedreven wordt door een Large Language Model (LLM).

De stakeholder heeft aangegeven dat er niet altijd een docent is die in staat is om op elk moment alle studenten direct te ondersteunen wanneer ondersteuning wordt gevraagd. In veel gevallen is het simpelweg niet haalbaar om tegelijkertijd alle vragen binnen een groep studenten te beantwoorden, vooral wanneer sprake is van een grote groep met complexe vraagstukken.

Om deze knelpunten op te vangen, wordt in dit project, verder gebouwd op het ART-project, een virtuele avatar ontwikkeld die studenten op een interactieve en intelligente manier ondersteuning biedt wanneer directe docentbegeleiding niet beschikbaar is.

## 1.2 Doelstelling

Het doel van het project is om een virtuele avatar te ontwikkelen voor het DI-lab, die uitbreidt op de eerdere resultaten van het ART-project. Dit project is dan ook een vervolgproject, waarbij de opgedane kennis uit het ART-project verder wordt uitgebreid en toegepast in een nieuwe, meer interactieve en intelligente toepassing.

De virtuele avatar krijgt als primaire functie het begeleiden van studenten door middel van interactieve vragen te beantwoorden en te stellen. Door gerichte vragen te stellen, kan de avatar de student ondersteunen bij het leerproces, bijvoorbeeld door leerdoelen in kaart te brengen, voortgang te monitoren of gerichte feedback te geven op basis van de antwoorden van de student.

Daarnaast is het een belangrijk doel dat de avatar in staat is om zichzelf continu te verbeteren door gebruik te maken van feedbackmechanismen[2]. Door middel van Continual Learning kan de avatar leren van eerdere interacties met studenten en docenten. Op deze manier wordt de avatar steeds slimmer en beter in staat om gepersonaliseerde begeleiding te geven.

Om de interactie met de avatar zo realistisch en betrokken mogelijk te maken, wordt er gebruikgemaakt van Augmented Reality (AR)[13]. Door AR toe te passen, kan de avatar in een fysieke leeromgeving van de student worden geplaatst, waardoor er een meer natuurlijke en intuïtieve interactie ontstaat. Hierdoor ontstaat er een meer persoonlijke en realistische leerervaring, waarbij de virtuele begeleider letterlijk onderdeel wordt van de leeromgeving.

Met dit project wordt niet alleen verder gebouwd op een eerder project, maar wordt er ook een innovatieve oplossing gecreëerd die kan bijdragen aan een effectiever, efficiënter en meer gepersonaliseerd leerproces binnen het DI-lab.

# 2 Methode

Voor dit project is gewerkt met een ontwerpergerichte aanpak volgens de principes van Design Science Research (DSR) [7]. Binnen deze methode staat het ontwerpen, bouwen en testen van een innovatief artefact centraal. Het doel van dit project was het verbeteren van begeleiding in het onderwijs, zonder tussenkomst van een docent, door middel van een virtuele avatar.

Het project is uitgevoerd in drie fasen:

- De onderzoeksfase, waarin de context, gebruikersbehoeften en technische mogelijkheden zijn onderzocht;
- De prototypefase, waarin op iteratieve wijze een werkend softwareprototype van de avatar is ontwikkeld;
- De communicatie- of implementatiefase, waarin het prototype kan communiceren door middel van tekst en spraak. Ook kan de avatar worden gebruikt op telefoon en in AR.

Elke fase gebruikte passende methodes zoals gebruikersonderzoek, vergelijkend onderzoek, het ontwerpen en bouwen van prototypes, en het testen van functionaliteiten. Hieronder worden

de drie fasen toegelicht.

## 2.1 Fase 1 – Onderzoek

De eerste fase richtte zich op het verzamelen van kennis om goede keuzes te maken voor de ontwikkeling van de avatar. Vier hoofdonderwerpen zijn onderzocht: ontwikkelloftware (engines), avatarcreatie, augmented reality (AR) en continual learning. Daarnaast is een vragenlijst uitgezet onder studenten.

### Vergelijking onderzoek engines

Er is onderzoek gedaan naar vier engines: Unity, Godot, Unreal Engine en Evergine. Deze zijn vergeleken op basis van gebruiksvriendelijkheid, prestaties, prijs, AR-ondersteuning en documentatie om te bepalen welke het beste past binnen het project. Informatie is verzameld via Google en officiële documentatie van de engines.[14, 3, 6, 4]

### Avatarcreatie

Ook is er onderzoek gedaan naar avatar creators. Blender en Ready Player Me zijn vergeleken op basis van creativiteit, complexiteit en integratie met Unity. Hiervoor is gebruikgemaakt van officiële documentatie en grijze literatuur.[9, 5]

### Augmented Reality

De AR-mogelijkheden van Unity zijn onderzocht met aandacht voor AR Foundation en ondersteuning op mobiele telefoons. Hiervoor is gebruikgemaakt van Unity-documentatie, engineonderzoek, online bronnen en gesprekken met onderzoekers van het DI-lab.[13]

### Continual Learning

Er is gekeken hoe taalmodellen kunnen blijven leren. Hiervoor is grijze literatuur geraadpleegd en zijn gesprekken gevoerd met onderzoekers van het DI-lab tijdens geplande en informele momenten. [2]

### Design onderzoek

Met een vragenlijst zijn studentvoorgekeuren verzameld over stijl, perspectief en functionaliteiten van de avatar. Deze is door 40 studenten ingevuld tijdens een wekelijkse stand-up. Vragen gingen onder andere over 2D of 3D stijl, volledige lichaamsweergave of alleen een torso, en gewenste functies zoals small talk of animaties.

## 2.2 Fase 2 – Prototype ontwikkeling

Op basis van de onderzoeksresultaten is gestart met het ontwikkelen van een werkend prototype.

### Unity structuur

Een nieuw Unity-project is opgezet met scènes voor het startscherm, de AR-scène, instellingen en de avatarinterface.[14]

### Avatar implementatie

De avatar is toegevoegd via Ready Player Me en aangepast op basis van de resultaten van de vragenlijst. De standaardinstellingen zijn aangepast aan de voorkeuren van gebruikers en stakeholders. [9]

### Achtergrond

In de vragenlijst werden drie achtergrondopties gegeven: een skybox, een muur met afbeelding of een 3D-omgeving. De meest gekozen optie is geïmplementeerd in het prototype.

### Gebruikersinterface

UX-elementen zoals belichting, camerastandpunten en schermindeling zijn ontworpen met ondersteuning van het DI-lab, via wekelijkse overleggen.

## 2.3 Fase 3 – Communicatie en Implementatie

De laatste fase richtte zich op communicatie tussen gebruiker en avatar, en op het functioneel maken van de applicatie voor mobiel en AR.

### LLM-integratie

Er is gezocht naar geschikte lokale taalmodellen via grijze literatuur en input van het DI-lab. Eerst is geprobeerd Neocortex in Unity te gebruiken, maar na feedback van stakeholders is gekozen voor communicatie via een API met een lokaal taalmodel (Ollama LLaMA 3). [10, 1]

### Test-API

Een test-API is opgezet om dummydata te versturen en ontvangen tussen Unity en de backend.

#### **Text-to-Text**

Vervolgens is een API gemaakt met Flask voor TextToText-functionaliteit. Hierbij is een script genaamd TextToTextClient in Unity ontwikkeld dat de API aanspreekt en antwoorden toont. Het gekozen taalmodel is lokaal geïnstalleerd en met een payload gekoppeld aan het script. [12]

#### **Speech-to-Text**

Na de TextToText-module is SpeechToText geïmplementeerd. Hierbij wordt spraak omgezet in tekst met Whisper van OpenAI. Dit is verwerkt in zowel de API als Unity met het script SpeechToTextClient.[11]

#### **Mobiel en AR**

Het invoersysteem is aangepast voor compatibiliteit met meerdere apparaten. Het canvas is geoptimaliseerd en een AR-camera toegevoegd in de AR-scène.

#### **Build en testen**

De applicatie is gebouwd en getest voor desktop en mobiel. Hiervoor zijn verschillende build-profielen aangemaakt, en is getest op stabiliteit en gebruiksvriendelijkheid.

### **3 Resultaten**

Het resultaat van dit project is een werkend softwareprototype van een virtuele avatar genaamd DI.ON.NE. Dit artefact betreft een interactieve softwareapplicatie die is ontwikkeld in Unity en functioneert op desktop, mobiel en in Augmented Reality (AR). De applicatie stelt gebruikers in staat om vragen te stellen via tekst of spraak, waarop de avatar een antwoord genereert met behulp van een lokaal taalmodel. Het doel van dit artefact is het bieden van digitale ondersteuning aan studenten binnen het onderwijs.

#### **3.1 Avatarontwerp en gebruikersvoorkeuren**

De avatar is visueel vormgegeven op basis van Ready Player Me en aangepast aan de voorkeuren van studenten, verzameld via een vragenlijst. Gebruikers gaven de voorkeur aan een 3D-avatar met alleen een bovenlichaam (torso), wat een goede balans biedt tussen visuele aanwezigheid en technische prestaties. Ook werd gekozen voor een neutrale, statische achtergrond in plaats van een volledige 3D-omgeving of een skybox. De stijl van de avatar is zo aangepast dat deze past binnen de voorkeuren voor uitstraling en interactie, met lichte animaties en een vriendelijke gezichtsuitdrukking.

#### **3.2 Text-to-Text communicatie**

De TextToText-functionaliteit stelt gebruikers in staat om via een invoerveld in Unity vragen te typen. Deze input wordt via een lokaal gehoste API verstuurd naar een taalmodel (LLaMA 3 via Ollama), gehost op dezelfde machine. De API is opgezet met Flask, inclusief CORS-ondersteuning voor communicatie met Unity. Binnen Unity is een script genaamd TextToTextClient ontwikkeld dat de input verwerkt, de API aanroeft en het antwoord weergeeft in het chatvenster. Dit antwoord wordt vervolgens getoond in de interface van de avatar, waarmee een natuurlijke conversatie ontstaat. [12, 1, 10]

#### **3.3 Speech-to-Text integratie**

Om spraakinteractie mogelijk te maken is gebruikgemaakt van Whisper, een open-source spraakherkenningsmodel van OpenAI. Gebruikers kunnen een spraakopname starten door op een knop te drukken of de ‘R’-toets ingedrukt te houden. De audio wordt opgenomen als een .wav-bestand en verzonden naar een API-endpoint, waar Whisper de audio omzet naar tekst. Deze tekst wordt vervolgens doorgestuurd naar dezelfde TextToText-module als reguliere tekstinvoer. De bijbehorende functionaliteit is geïmplementeerd in het Unity-script SpeechToTextClient.[11]

#### **3.4 Platformonafhankelijkheid en AR-functionaliteit**

Een belangrijk onderdeel van het project was de wens om de applicatie functioneel te maken op verschillende platformen. De interface is daarom ontworpen met een responsive canvas dat zich automatisch aanpast aan verschillende schermformaten. Voor mobiele ondersteuning is een Android-build uitgevoerd en getest. Daarnaast is in Unity een AR-scène opgezet met

behulp van AR Foundation, waarmee de avatar in een echte omgeving geprojecteerd kan worden via de camera van een mobiel apparaat.[13]

### 3.5 Privacy en lokale verwerking

Een centrale ontwerpkeuze was om het systeem volledig lokaal te laten draaien. Dit voorkomt dat gevoelige gebruikersdata extern wordt verwerkt, wat aansluit bij de wensen en eisen van het onderwijsveld. Zowel de taalmodellen als de API's draaien lokaal, wat de privacy verhoogt en het systeem onafhankelijk maakt van een internetverbinding. Deze keuze werd mede ingegeven door gesprekken met onderzoekers van het DI-lab en de opdrachtgever. [10]

### 3.6 Testresultaten en gebruikerservaring

Tijdens gebruikstests werd het prototype positief beoordeeld. Gebruikers rapporteerden een snelle verwerking van vragen (binnen enkele seconden) en een duidelijke interface. Vooral de lage drempel om een vraag te stellen werd gewaardeerd. De avatar kwam betrouwbaar en vriendelijk over, wat de bereidheid tot interactie vergrootte. Zowel de spraak- als tekstfunctionaliteit functioneerde stabiel en consistent tijdens tests op desktop.

### 3.7 Samenvatting van het artefact

Het eindproduct is een modulaire, lokaal draaiende avatarapplicatie met ondersteuning voor tekst- en spraakinteractie. Het artefact is gebaseerd op inzichten uit gebruikersonderzoek, vergelijkend technisch onderzoek en iteratief prototyping. De applicatie vormt een solide basis voor verdere uitbreiding, zoals integratie van text-to-speech, mimiek, emotieherkenning of koppeling aan onderwijspecifieke kennisbanken. Hiermee is voldaan aan de kernvereisten: een toegankelijke, platformonafhankelijke, privacyvriendelijke virtuele begeleider voor het onderwijs.

## 4 Discussie

Het project heeft geleid tot een functionerend prototype van een digitale avatar die studenten kan ondersteunen in het onderwijs. Er zijn verschillende factoren geweest die positief hebben bijgedragen aan dit resultaat, maar ook enkele die het proces hebben vertraagd of beperkt.

Een belangrijke positieve factor was de keuze voor een iteratieve werkwijze volgens het Design Science Research-model.[7] Door het werk op te delen in duidelijke fases — onderzoek, prototypeontwikkeling en implementatie — ontstond er structuur en ruimte voor reflectie en verbetering per fase. Vooral de onderzoeksfase droeg sterk bij aan het maken van goed onderbouwde ontwerpkeuzes. Het vergelijken van engines, avatarplatforms en AR-mogelijkheden voorkwam verspilling van tijd en middelen in latere stadia. Ook de vragenlijst onder studenten leverde waardevolle inzichten op in gebruikersvoorkeuren, waardoor het ontwerp beter kon aansluiten op de doelgroep.

Daarnaast werkte de ondersteuning vanuit het DI-lab stimulerend. De combinatie van technische begeleiding en regelmatige feedbackmomenten droeg bij aan het versneld oplossen van obstakels in de ontwikkelfase. Ook de keuze om te werken met Unity en Ready Player Me bleek effectief: beide tools sloten goed aan bij de gewenste functionaliteit en boden voldoende ondersteuning via documentatie en community.

Tegelijkertijd waren er ook beperkende factoren. Een eerste beperking was de beperkte tijd om een volledig volwassen product op te leveren. Hoewel het prototype functioneert, is het nog geen eindproduct dat zonder begeleiding grootschalig inzetbaar is. Veel tijd ging verloren aan het verkennen van alternatieve routes — zoals het initiële experiment met Neocortex — die later moesten worden verlaten. Dit had voorkomen kunnen worden door vooraf duidelijke eisen en randvoorwaarden te definiëren met stakeholders.

Een tweede beperkende factor was het gebrek aan ervaring met back-endontwikkeling en taalmodellen. Hierdoor kostte het opzetten van een werkende API met Whisper en LLaMA relatief veel tijd. Ook het testen op meerdere platformen (desktop, mobiel, AR) leverde extra technische complexiteit op.

Tot slot bleek dat het combineren van spraakfunctionaliteit en AI in een Unity-applicatie veel technische kennis vereist. Hoewel het prototype deze functies bevat, is verdere optimalisatie nodig om een robuust, schaalbaar systeem te realiseren.

## 5 Conclusie

Het oorspronkelijke doel van dit project was het ontwikkelen van een digitale avatar die studenten kan ondersteunen bij het beantwoorden van vragen, zonder tussenkomst van een docent. Dit doel is grotendeels bereikt. Er is een functionerend softwareprototype ontwikkeld dat tekstuele en gesproken interactie mogelijk maakt, gebruikmakend van een lokaal taalmodel. De avatar werkt binnen een Unity-applicatie en is inzetbaar op zowel desktop als mobiele apparaten, inclusief ondersteuning voor Augmented Reality.

Binnen het project zijn meerdere artefacten gerealiseerd. Ten eerste is er een werkende avatar ontworpen en geïmplementeerd met behulp van Ready Player Me [9], afgestemd op voorkeuren van studenten die via een enquête zijn verzameld. Daarnaast zijn er API's ontwikkeld voor zowel Text-to-Text (LLM-integratie via LLaMA 3 en Ollama)[1, 10] als Speech-to-Text (via Whisper)[11]. Deze API's maken communicatie mogelijk tussen Unity en de taalmodellen. Ook is een functionele gebruikersinterface ontwikkeld en zijn Unity-scènes opgezet voor verschillende gebruikssituaties, zoals AR en mobiel gebruik.

De meeste randvoorwaarden en technische vereisten zijn vervuld. De applicatie kan vragen ontvangen via tekst of spraak, verwerken via een lokaal AI-model, en antwoorden tonen in de interface. Daarnaast is de toepassing getest op verschillende platformen. Hoewel de applicatie functioneert, is het prototype nog niet geschikt voor grootschalige implementatie in het onderwijs. Er ontbreken uitgebreide gebruikers- en stresstests, en er is nog geen koppeling met leermanagementsystemen (zoals Moodle of Canvas).

Op basis van de resultaten kunnen enkele aanbevelingen worden gedaan. Richting de praktijk wordt aangeraden om het prototype in een kleinschalige onderwijssetting te testen met echte gebruikers. Dit levert inzichten op in gebruiksvriendelijkheid, effectiviteit en impact op het leerproces. Ook is het wenselijk om functies als personalisatie, feedbacklogboeken of progressietracking toe te voegen.

Voor vervolgonderzoek wordt aanbevolen om te verkennen hoe AI-avatars op lange termijn kunnen samenwerken met menselijke docenten, en hoe ze kunnen worden geïntegreerd in bestaande onderwijsystemen. Ook is het zinvol om te onderzoeken hoe de gebruikte AI-modellen beter kunnen worden aangepast aan educatieve doeleinden en specifieke domeinkennis.

## Verwijzingen

- [1] Meta AI. Llama 3. <https://ai.meta.com/llama>, 2024.
- [2] Matthijs Delange and Marc en Parisot Sarah en Jia Xu en Leonards Aleš en Slabaugh Greg en Tuytelaars Tinne Aljundi, Rahaf en Masana. A continual learning survey: Defying forgetting in classification tasks. <https://arxiv.org/abs/1909.08383>, 2021.
- [3] Godot Engine. Godot engine. <https://godotengine.org>, 2025.
- [4] Evergine. Evergine 3d engine. <https://evergine.com>, 2025.
- [5] Blender Foundation. Blender. <https://www.blender.org>, 2025.
- [6] Epic Games. Unreal engine. <https://www.unrealengine.com>, 2025.
- [7] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. Design science in information systems research. *MIS Quarterly*, 28(1):75–105, 2004.
- [8] Zuyd Hogeschool. Di-lab. <https://www.zuyd.nl/onderzoek/lektoraten/data-intelligence>, 2024.
- [9] Ready Player Me. Ready player me: Avatar platform for the metaverse. <https://readyplayer.me>, 2025.
- [10] Ollama. Ollama: Run llms locally. <https://ollama.com>, 2025.
- [11] OpenAI. Whisper: Speech recognition model. <https://openai.com/research/whisper>, 2022.
- [12] Armin Ronacher. Flask: Python micro web framework. <https://flask.palletsprojects.com>, 2025.
- [13] Unity Technologies. Ar foundation. <https://docs.unity3d.com/Packages/com.unity.xr.arfoundation>, 2025.
- [14] Unity Technologies. Unity real-time development platform. <https://unity.com>, 2025.

# Bijlagen

Vraag:	Antwoord:	Vraag 1 optie A: Hoofd	Vraag 2 optie A: 3D	Vraag 3 optie A: Small Talk	Eigen idee?
Vraag 1: Voorkeur Avatar Lichaam	B				X
Vraag 2: Voorkeur Avatar Stijl	B / C	D			X
Vraag 3: Voorkeur Avatar Functies	C				X
		Vraag 1 optie B: torso	Vraag 2 optie B: 2D	Vraag 3 optie B: Idle Animation	
		Vraag 1 optie C: Heel Lichaam	Vraag 2 optie C: Cartoon	Vraag 3 optie C: Emotie Utting	
				Vraag 3 optie D: Gamification	
Naam:	Antwoord 1: Antwoord 2: Antwoord 3:	Antwoord 1 Waarden:	Antwoord 2 Waarden:	Antwoord 3 Waarden:	Eigen idee?
Capivara	B B B / C	A A D	Realistisch Heel lichaam heeft een goede afstand	Realistischer SD kan erg weleven	Mannelijke en vrouwelijke avatars
?	C	D	Heel lichaam voegt geen waarde toe	Less is more	Laat de gebruiker de Avatar samenstellen
Quentin S	A	C	D	Gamification zorgt voor motivatie	
Kenneth	A	B	D	Meer motivatie	
John	B	B	D	Meer interactieve functie	
Ramon	B	A	D	Professioneler	
Stijn S	B	A	D	Motiverend	
Max V + Gianfranco K	C	A	D	Steriler en realistischer	
?	B	C	D	Lijk realistisch	
Golan	C	A	D	Houdt meer aansprekelijk, lichaam is niet leuk.	
Jesper K	B	A	D	Leuke en aansprekelijk, maar niet te dichtbij	
?	C	A	D	Realistisch	
Jonah	B	B	D	Met de handen is het heel verschillend, maar niet te dichtbij	
Fenkie + Jasmina	B	C	D	Ziet er fijner uit	
Danny	C	C	D	Gaming elementen zijn leuk	
?	B	C	D	Cartoons zijn meer aansprekbaar	
Tou + Casper	B	A	D	Helpt beter met leren	
?	B	A	D	Avatar moet meer interactie hebben	
Timothy + Tony	C	A	D	Avatar kan muziek spelen	
Justin S	C	A	D	Avatar zelf kunnen maken	
?	B	B	C	Punten kunnen wegskalen	
Shelly + Eunice	C	A	D	Op een positieve manier en de avatar laten praten	
Sean D	B	A	D	Het blikt een robot	
Finn v K	B	C	A / C	Een winkel om de punten te spenderen op rewards	
Daan + Rinse	B	A	D	Avatar moet meer interactie hebben	
Tijn M	B	A	A / D	Om meer te kunnen doen	
Joyren T	A	A	A / B / C	Uitbreiden tot een toernament	
				Om meer leaven aan de avatar te geven	
Meest gekozen:	B A D				

Figuur 1: Resultaat van het onderzoek: Avatar design