

# Scholieren met dyslexie van de derde graad middelbaar onderwijs ondersteunen bij het begrijpend lezen van wetenschappelijke artikelen via gepersonaliseerde ATS.

Dylan Cluyse

dylan.cluyse@student.hogent.be

Promotor: Lena De Mol

Co-promotor: Johan Decorte (Hogeschool Gent); Jana Van Damme (Hogeschool Gent)

Hogeschool Gent, Valentin Vaerwyckweg 1, 9000 Gent

## Samenvatting

Ingewikkelde woordenschat en zinsbouw hinderen scholieren met dyslexie in de derde graad van het middelbaar onderwijs bij het begrijpend lezen van wetenschappelijke artikelen. Gepersonaliseerde manual text simplification (MTS) helpt deze scholieren bij hun leesbegrip. Daarnaast kan automatic text simplification (ATS) dit proces automatiseren om de werkdruk bij leraren en scholieren te verminderen. Dit onderzoek achterhaalt met welke technologische en logopedische aspecten AI-ontwikkelaars rekening moeten houden bij de ontwikkeling van een AI-toepassing voor geautomatiseerde en gepersonaliseerde tekstvereenvoudiging. Hiervoor stelt het onderzoek de volgende onderzoeksvraag op: "Hoe kan een wetenschappelijk artikel automatisch worden vereenvoudigd, gericht op de unieke noden van scholieren met dyslexie in de derde graad middelbaar onderwijs?". Een requirementsanalyse achterhaalt de benodigde functionaliteiten om gepersonaliseerde ATS mogelijk te maken. Vervolgens wijst de vergelijkende studie uit welk taalmodel ontwikkelaars kunnen inzetten om de taak van gepersonaliseerde en geautomatiseerde tekstvereenvoudiging mogelijk te maken. De requirementsanalyse wijst uit dat toepassingen om wetenschappelijke artikelen te vereenvoudigen, gemaakt zijn voor een centrale doelgroep en geen rekening houden met de unieke noden van een scholier met dyslexie in de derde graad middelbaar onderwijs. Toepassingen voor gepersonaliseerde ATS zijn mogelijk, maar ontwikkelaars moeten meer inzetten op de unieke noden van deze scholieren.

Keuzerichting: AI en data engineering

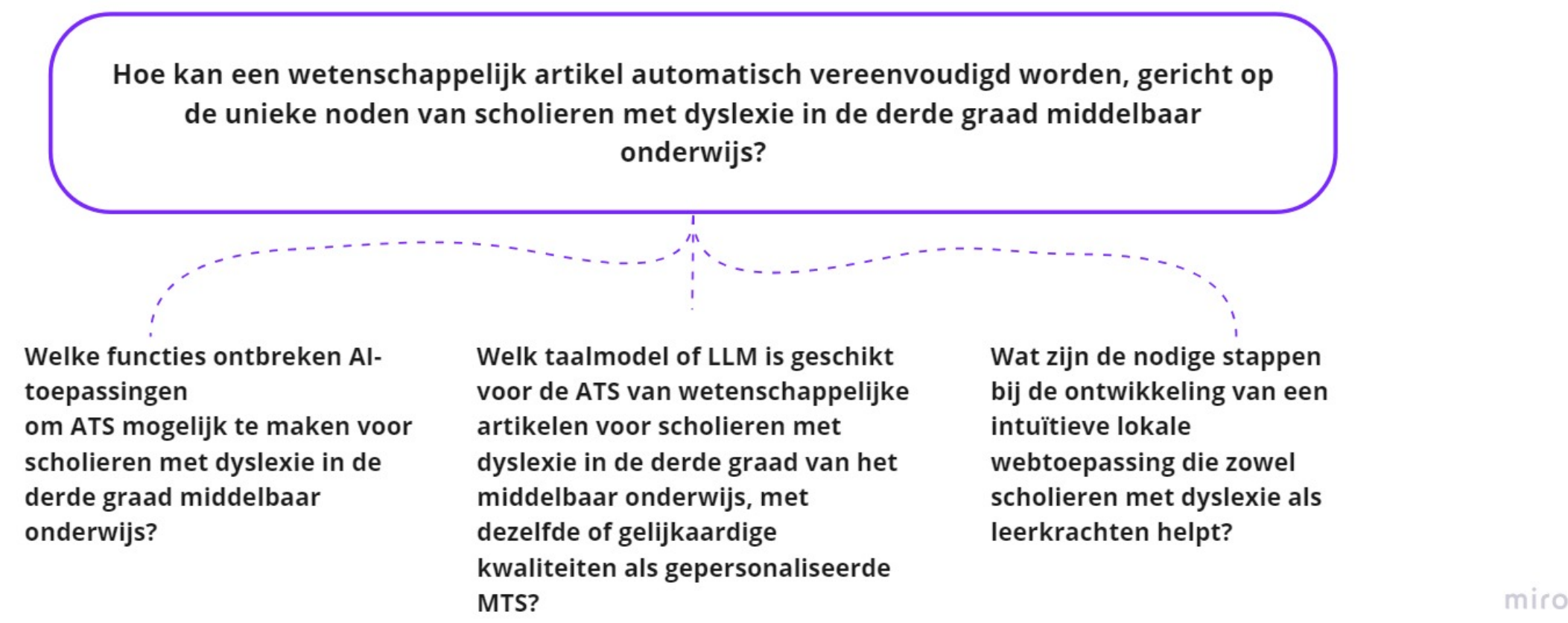
Sleutelwoorden: Gepersonaliseerde tekstvereenvoudiging; dyslexie; natuurlijke taalverwerking

Broncode: <https://github.com/dylancluyse/bachelorproef-nlp-tekstvereenvoudiging>

## 1. Introductie

Ik wkgndeleie worechnsadot en zinsdouw hdirnnen scholieren met dlyix-see in de drdee graad middeldaar onberwijs bij het lezen van wkiclnphtja-pepe ateirkeln.

Vindt u deze zin moeilijk te lezen? Dit kan voor scholieren met dyslexie dagelijkse kost zijn bij het begrijpend lezen. Stoort het vakjargon en veel informatie in een compact formaat u bij het begrijpend lezen van nieuw wetenschappelijk onderzoek? Dit is niet abnormaal, want het begrijpend lezen van wetenschappelijke artikelen vraagt een alsmaar grotere geletterdheid van de lezer. Toch kunnen leerkrachten deze hindernis voor de studenten verwerpen, door teksten met *manual text simplification* (MTS) te vereenvoudigen. Om de werkdruk in het onderwijs tegen te gaan, moet dit proces ook geautomatiseerd worden met *automatic text simplification* (ATS). Figuur 1 illustreert de centrale onderzoeksvraag, gekoppeld aan de deelvragen die het onderzoek moet beantwoorden.



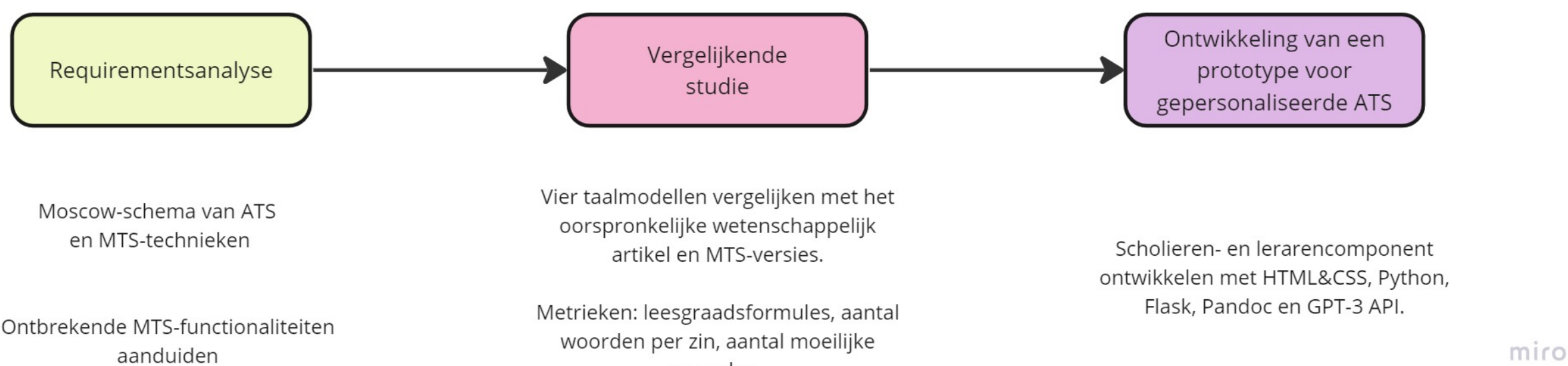
Figuur 1: Centraal de onderzoeksvraag.

## 2. Onderzoeksmethoden

Om een werkwijze te achterhalen voor het ontwikkelen van dergelijk toepassing, voert het onderzoek drie fasen uit weergegeven op figuur 2.

De requirementsanalyse toetst de huidige toepassingen af op hun MTS-functionaliteiten.

De vergelijkende studie van taalmodellen gebeurt met machinale en menselijke beoordeling. Het onderzoekt laat de machine op basis van leesbaarheidsscores en het aantal (complexe/ lange) woorden per zin. Daarnaast maakt het onderzoek gebruik van referentieteksten vereenvoudigd door leerkrachten en scholieren om de gelijkenissen en gebreken te achterhalen van ATS-teksten met MTS-teksten.



Figuur 2: De gebruikte onderzoeksmethoden. In het vetgedrukt staan de onderzoeksvragen beschreven.

## 3. Resultaten

Scholieren kunnen de tekst aanpassen. Daarnaast kunnen leerkrachten een vereenvoudigde versie van een wetenschappelijk artikel laten maken met het prototype. Hier kan de leerkracht parameters meegeven, waaronder de opmaakopties en de vereenvoudigingen die de tekst moet aannemen.

## 4. Conclusies

Uit de requirementsanalyse blijkt dat bestaande online tools en softwaretoepassingen voor ATS onvoldoende gepersonaliseerde functionaliteiten bieden. Ze missen opmaakopties voor scholieren met dyslexie. Recente technologieën bieden wel tekstvereenvoudigingsopties, maar vereisen informaticakennis die de meeste scholieren en leraren niet hebben. Een vergelijkende studie toont aan dat het GPT-3 taalmodel geschikt is voor LS en SS-technieken, terwijl andere taalmodellen minder coherente tekst genereren. Het prototype voor gepersonaliseerde ATS maakt gebruik van AI en NLP-technologieën zoals PDFMiner, de GPT-3 API en Pandoc. Hoewel het prototype nog niet aan alle functionaliteiten voldoet, kunnen ontwikkelaars met de gebruikte softwarepakketten een volledig afgewerkte toepassing ontwikkelen. Figuur 3 toont de verdere onderzoeksonderwerpen aan.



Figuur 3