

Algemene schetsing

- We zitten allemaal terug in de laatste graad van het middelbaar onderwijs.
- Om bewust te worden van wetenschappelijk onderzoek, krijgen wij tijdens onze STEM-vakken een wetenschappelijk artikel voorgeschoteld.
- Een opdracht begrijpend lezen kan ons doen inzien waarom wetenschappelijk onderzoek wordt gedaan, maar ook om ons actueel te houden van de laatste wetenschappelijke inzichten.
- Deze structuur herkennen we ongetwijfeld, maar wat als ik nu een aanpassing laat maken aan de tekst.

Demo 1

- Wat ik jullie nu toon is geen optisch fenomeen, maar wel iets waar scholieren in het middelbaar onderwijs mee kunnen te maken hebben.
- Deze simulatie benadrukt hoe moeilijk het begrijpend lezen van deze artikelen kan worden voor scholieren met dyslexie.
- Onderzoeken wijzen uit hoe ontwikkelaars de digitale weergave van deze teksten kan verbeteren. Zo kan het lettertype naar een sans-serif font aangepast worden, het lettertype kunnen we opschalen naar 15pt en de achtergrondkleur kan aangepast worden om deze scholieren een aangename leeservaring aan te bieden.
- Toch blijft de tekst in het wetenschappelijk artikel een obstakel voor alle scholieren in het middelbaar onderwijs. Onderzoeken wijzen uit dat de benodigde geletterdheid van deze teksten alsmaar stijgt, waardoor een steeds groter percentage de grip op deze materie verliest.
- Bestaande toepassingen de tekst op een nieuwe manier weergeven, maar het artikel beschikt nog steeds over wetenschappelijk jargon, lange zinnen die over meerdere lijnen kunnen spreiden, meerlettergreperige woorden of complexe zinsyntax.
- Hiervoor bestaat er wel al een oplossing. Zo kunnen leerkrachten wetenschappelijke artikelen vereenvoudigen op maat van deze scholieren. Zij kunnen de tekstinhoud van een wetenschappelijk artikel overnemen en in een Word-document de opmaak aanpassen, maar ook de tekst herschrijven gebaseerd op de achtergrondkennis van de leerlingen.

Demo 2

Er zijn drie algemene MTS-technieken:

1. *Lexical simplification* kunnen leerkrachten doen door de moeilijke woorden in de tekst te vervangen door een eenvoudiger synoniem. Als er geen eenvoudiger synoniem bestaat, kan de leerkracht de zin herschrijven zodat die over een in-line betekenis beschikt. Ondersteunend hierop kan de leerkracht ook een woordenlijst schrijven dat de moeilijke woorden kort uitlegt. Moeilijke woordenschat kan vervangen worden. Echter moeten we rekening houden met de doelgroep. Reeds gekend jargon behoort nog steeds tot de te kennen leerstof en deze moeten blijven staan. Bij MTS komt een didactisch aspect bij kijken. Zo moet dergelijk prototype enkel ondersteuning bieden en geen vervangende tekst genereren.
 2. Daarnaast is er *Syntactic simplification* door zinnen te verkorten, passief naar actief omvormen, zinnen met verbindingswoorden opbreken. Het doel is om kortere zinnen te verkrijgen met een gelijke semantiek. Zo kunnen scholieren dezelfde tekst lezen, maar beter volgen.
 3. Structurele aanpassingen, door de tekst in een ander formaat te schrijven. Bijvoorbeeld als opsomming of in tabelvorm. Samenvatten verkort de tekst, maar is niet noodzakelijk een tekstvereenvoudiging. Zo kunnen moeilijke zinsconstructies of woorden nog steeds in een samenvatting terug te vinden zijn.
- Hoewel er geen gekend percentage is over het aantal leerkrachten dit doen, toch is dit een techniek die een bewezen effect heeft voor zowel scholieren met dyslexie als scholieren zonder dyslexie.
 - Nu heb ik jullie de vier technieken getoond om een abstract te vereenvoudigen, maar hoe zit het dan met een wetenschappelijk artikel van minstens 10000 woorden?
 - Daarnaast houdt de huidige vereenvoudiging enkel rekening met één doelgroep.
 - Meerdere versies maken is een moeizame taak binnen het onderwijs, dat zich nu al op een kraakpunt bevindt.

- Dergelijk toepassingen kunnen alle factoren binnen het onderwijs beïnvloeden.
- Dit onderzoek achterhaalt hoe ontwikkelaars dergelijk toepassingen kunnen ontwikkelen.

Requirementsanalyse

- Allereerst toetst de requirementsanalyse de MTS-technieken bij bestaande software af.
- Zo toetst de requirementsanalyse geen proof-of-concepts of andere prototypen af.
- De tools bestaan uit drie erkende softwarepakketten, namelijk Kurzweil, Sprint en Alinea.
- Daarnaast toetst het onderzoek vijf vrij beschikbare toepassingen af, namelijk Rewordify, Scispace, Simplish, Bing Chat en ChatGPT.
- Twee gepubliceerde wetenschappelijke artikelen dienen als voorbeeld om de functionaliteiten van de tools te achterhalen.
- Op basis van de MTS-technieken bouwt het onderzoek een Moscow-schema op. De must-haves omvatten alle nodige technieken en functionaliteiten om een wetenschappelijk artikel gepersonaliseerde te kunnen vereenvoudigen.
- Bijvoorbeeld het selecteren van een zin om deze syntactisch te vereenvoudigen, of het markeren van een woord om deze aan de woordenlijst toe te voegen.
- Opmaakopties in de webtoepassing en in het vereenvoudigde document toetst het onderzoek ook af. Zo kan het onderzoek achterhalen of deze toepassingen rekening houden met scholieren met dyslexie in de derde graad van het middelbaar onderwijs.
- Op basis van de vereenvoudigde documenten, kan het onderzoek de ontbrekende functionaliteiten voor gepersonaliseerde ATS achterhalen. Daarnaast staat het onderzoek stil bij de ontbrekende MTS-technieken in huidige toepassingen.

Bestaande toepassingen

- Van deze vijf staan stil bij Kurzweil, Sprint en Alinea. Deze pakketten zijn op maat gemaakt voor scholieren met dyslexie door luistersoftware aan te reiken in een aanpasbare omgeving. Zo kunnen scholieren de lettertype- en spatiëring aanpassen naar hun wens.
- Deze pakketten bieden geen MTS-technieken aan, buiten het genereren van woordenlijsten. De tekst blijft zoals voordien.

Online toepassingen

- Online vinden we wel weer andere toepassingen terug. Deze beschikken over meer lexical en syntactic simplification technieken, maar ontbreken structurele aanpassingen en ondersteunende visuele opties voor scholieren met dyslexie.
- Op basis van deze toepassingen kunnen de must-have technieken en functionaliteiten opgehaald worden. Daarnaast zijn de beschikbare tools eerder gericht op het samenvatten van teksten, wat niet noodzakelijk bijdraagt tot een vereenvoudiging.

Vergelijkende studie

- Taalmodellen genoeg ! Of toch niet?
 - Welke taalmodellen maken gebruik van vereenvoudiging?
 - Welke taalmodellen zijn getraind op het vereenvoudigen van wetenschappelijke artikelen, specifiek voor scholieren van deze leeftijdsgroep?
 - Welke taalmodellen zijn personaliseerbaar?
 - Welke taalmodellen beschikken over zoveel parameters dat dit er niet toe doet?

- Met de nodige functionaliteiten, moet het onderzoek een geschikt taalmodel voor personaliseerbare ATS kunnen achterhalen. Zo moet het taalmodel lange artikelen kunnen verwerken, maar rekening houden met de persoonlijke noden van de eindgebruiker.
- Als referentiemateriaal werden er aan vier mensen, waarvan twee leerkrachten en twee leerlingen, gevraagd om één van de wetenschappelijke artikelen handmatig te vereenvoudigen. Deze vereenvoudiging deden ze door middel van een referentietekst.
- De vergelijkende studie neemt vier taalmodellen op. T1 tot T3 komen uit HuggingFace en zijn taalmodellen specifiek getraind om wetenschappelijke artikelen te kunnen vereenvoudigen of samenvatten.
- T4 is GPT-3 ofwel het achterliggende taalmodel van Bing Chat en ChatGPT dat in de requirementsanalyse goede scores kon behalen. Omdat GPT-3 een prompt-gebaseerd taalmodel is, maakt de vergelijkende studie ook gebruik van drie verschillende prompts die elk een specifieke vereenvoudigingstechniek uitproberen.
 - De eerste prompt is een algemene vereenvoudiging zonder het meegeven van de doelgroep, gevolgd door een nadrukkelijk lexicale vereenvoudiging.
 - De derde prompt benadrukt een syntactische vereenvoudiging.

Scripts:

1. De vergelijking gebeurt op basis van de uitvoertekst van het taalmodel. Pythonscripts halen de tekst uit het oorspronkelijke wetenschappelijke artikel opgehaald met een OCR-techniek.
2. Vervolgens dienen de HF-taalmodellen een Engelse tekst te krijgen als invoer, daarmee vertaalt het Python script dit oorspronkelijk artikel naar het Engels.
3. Daarna worden alle zinnen uit de tekst gehaald met Spacy sentence embeddings, om deze vervolgens te laten itereren en aan het taalmodel te geven.
4. De uitvoer van deze tekstdocumenten slaat het script uiteindelijk op in een CSV-bestand waarbij statistieken rond de taalmodellen gevisualiseerd kunnen worden met Jupyter-notebooks.

Prototype

- Het prototype kan twee doelgroepen baten. Allereerst zijn er scholieren die een tool willen waar zij aanpassingen kunnen maken aan de tekst. Anderzijds leerkrachten die een gegenereerde tekst willen laten maken voor een specifieke doelgroep. Om scholieren met dyslexie extra te kunnen ondersteunen, biedt het prototype opmaakoptyes aan zowel binnen de site, als in het gegenereerde document.
- Temperature zorgt ervoor dat het taalmodel binnen de perken wordt gehouden. Zo wijkt het taalmodel niet van de oorspronkelijke context af. Daarnaast zorgt de top-p waarde voor een hoge woordfrequentie.

Scholierencomponent

- Het scholierencomponent combineert de kracht van het GPT-3 talenmodel, met de intuïteit uit toepassingen zoals Scispace.
- Scholieren kunnen namelijk tekst markeren en deze vervolgens laten vereenvoudigen.
- Bij de vereenvoudiging kunnen scholieren ook kiezen of deze tekst in doorlopende tekst moet; of per opsomming. De opsomming is handiger om een snel overzicht te hebben van de gemarkeerde tekst.
- Naast een ingebakken prompt, kunnen scholieren ook zelf een prompt schrijven. Zo kunnen ze een vraag stellen aan het taalmodel om hen te ondersteunen bij het begrijpen lezen.
- Om deze scholieren verder te ondersteunen, kan het prototype ook een woordenlijst opbouwen na een handmatige selectie van kernwoorden.
- Het prototype biedt ondersteuning aan, zonder didactische waarden van de scholier af te nemen.
- Kortom kunnen ontwikkelaars met vrij basiskennis van JS en Python de bestaande toepassingen recreëren.

- Scholieren in de derde graad moeten nog steeds nieuwe zinsstructuren en woorden aanleren.
- Daarom verwijdt het prototype geen nieuwe woordenschat of zinsstructuren in dit component.
- De oorspronkelijke tekst blijft staan. De woordenlijst zorgt ervoor dat leerlingen nieuwe woorden kunnen aanleren, ongeacht de moeilijkheidsgraad ervan.
- Als ontwikkelaar, pedagoog, leerkracht of logopedist is het moeilijk om de achtergrondkennis van deze scholieren te kunnen inschatten.
- Daarbij gaat ook een woordenschat te mee die zij moeten aanleren en niet door een machine mag vervangen worden. Daarom voegt het prototype woorden toe aan de huidige tekst.

Lerarencomponent

- Naast een ondersteunende tool, kan het prototype ook de werkdruk bij leerkrachten verminderen. Voordien moesten leerkrachten handmatig eerst de tekst doornemen.
- Enkel wordt er één prompt voor het volledige document gebruikt. Een volgende iteratie van het prototype zou per zin, per paragraaf of per hoofdstuk een prompt moeten krijgen.
- Alle tekstinhoud moet omgevormd worden naar een document dat scholieren kunnen terugkrijgen. Daardoor kunnen ontwikkelaars gebruikmaken van Pandoc. Alle vereenvoudigde tekstinhoud wordt naar een markdown-bestand uitgeschreven, die vervolgens naar een pdf of word document wordt omgezet. Met Pandoc kan het prototype de uitvoertekst in een personaliseerbaar formaat aanreiken.
- Wel is de opmaak beperkt bij Word / docx-bestanden. PDF's zijn opgebouwd met de Xelatex-engine en daardoor zijn alle opmaakoptyes om scholieren met dyslexie te helpen, parameteriseerbaar. Werken met pandoc laat de deur ook open om vereenvoudigde wetenschappelijke artikelen te genereren in epub (of eBook) formaat en eventueel slides genereren.

Conclusie

- Het prototype kan een ondersteunend middel aanbieden voor scholieren met dyslexie dankzij een carte blanche functionaliteit.
 - Scholieren kunnen eender welke aanpassing maken op de oorspronkelijke tekst, vooral gericht op de vereenvoudiging ervan.
- Hoewel Pandoc en GPT-3 een mooie start kunnen aanreiken voor gepersonaliseerde ATS, toch is het taalmodel niet 100% zeker voor welke doelgroep ATS moet uitgevoerd worden.
 - Daarnaast wordt enkel de tekstinhoud bekeken. Een verder onderzoek met GPT-4 moet aanduiden of dit taalmodel de doelgroep beter kan inschatten.
- Voor betere gepersonaliseerde ATS, kunnen volgende onderzoeken kijken naar *one-shot summaries*, door de eindgebruiker een paragraaf te laten schrijven waarop het taalmodel zich kan baseren, of door het opschalen naar GPT-4 dat over meer parameters beschikt.
- Daarnaast houdt het prototype enkel rekening met de tekstinhoud en niet met afbeeldingen.
- De volgende iteratie van het GPT-taalmodel is wel in staat om afbeeldingen te interpreteren en te genereren.
- Zo kan een volgend onderzoek stilstaan bij de capabiliteiten van GPT-4 bij het interpreteren van grafieken of gevisualiseerd cijfermateriaal.
- Daarnaast kan het mogelijk aanschouwelijkheid voor de gebruiker aanbieden door het taalmodel nieuwe afbeeldingen en schema's te laten genereren.