

## ONDERZOEKSVOORSTEL

# De kennis van de Info Support medewerkers opfrissen door het Info Support Guidance Framework te gamificeren met een quiz generator: een proof-of-concept.

Bachelorproef, 2022-2023

Manu Vleurick en Rayme Emin

E-mail: [manu.vleurick@student.hogent.be](mailto:manu.vleurick@student.hogent.be) en [rayme.emin@student.hogent.be](mailto:rayme.emin@student.hogent.be)

Co-promotor: Yoeri Van Damme (Info Support)

## Samenvatting

In dit onderzoek wordt een proof of concept in de vorm van een quiz generator gemaakt voor het Info Support Guidance Framework om te onderzoeken of dit effectief kan zijn in het opfrissen van het Guidance Framework. Het Guidance Framework is een framework waarin alle kennis die Info Support heeft opgedaan in zijn hele loopbaan staat beschreven. De quiz generator zou de medewerkers in staat kunnen stellen om op een laagdrempelige manier bij te blijven met de kennis beschreven in het framework. Een proof of concept van de quiz generator van de software oplossing wordt beschreven, waarbij flexibiliteit, gebruiksvriendelijkheid en uitbreidbaarheid van belang zijn. Er worden meerdere soorten vragen gegenereerd zoals meerkeuzevragen, waar/onwaar-vragen, invulvragen en combinatievragen. Voor elke soort vraag wordt er getest of dit relevant is tegenover de brontekst door middel van de BLUE score en manuele score-toereiking van een medewerker die het Guidance Framework gebruikt. Ook wordt er nagedacht over andere gamificatie methodes die er kunnen gebruikt worden om het Guidance Framework op te frissen.

**Keuzerichting:** AI & Data Engineering

**Sleutelwoorden:** NLP, Question Generation, Gamificatie, Guidance Framework

## Inhoudsopgave

1	Introductie . . . . .	1
2	State-of-the-art . . . . .	1
	2.1 Question generation . . . . .	1
	2.1.1 Meerkeuzevraag . . . . .	2
	2.1.2 Waar/onwaar-vraag . . . . .	2
	2.1.3 Invulvraag . . . . .	2
	2.1.4 Combinatievraag . . . . .	3
	2.2 Gamification . . . . .	3
3	Methodologie . . . . .	3
4	Verwacht resultaat, conclusie. . . . .	3

## 1. Introductie

Quizen worden gemaakt om kennis op te frissen dat je hebt geleerd. Maar wat als die kennis blijft uitbreiden en je niet zomaar altijd een nieuwe quiz kan maken? Dit is het probleem bij het Info Support Guidance Framework. Dit is een framework waar alle kennis die Info Support heeft verzameld tijdens hun carrière in de IT consultancy wereld wordt bijgehouden. Info Support wil door middel van gamificatie technieken de kennis van hun medewerkers opfrissen. Dit gebeurt door een Proof of Concept (PoC) van een automatische quiz generator die met Natural Language Processing (NLP) technieken verschillende soorten vragen kan genereren over de onderwerpen

beschreven in het Guidance Framework. De PoC zal relevante vragen moeten kunnen stellen om de essentie van de onderwerpen in het framework te toetsen. Er kunnen ook extra gamificatie technieken worden gebruikt naast de quiz.

## 2. State-of-the-art

### 2.1. Question generation

Het genereren van vragen met behulp van NLP-technieken is een uitdaging in de kunstmatige intelligentie. De technieken die gebruikt worden om vragen te genereren verschillen afhankelijk van het type vraag. Meerkeuzevragen en waar/onwaar-vragen zijn complexer om te genereren dan bijvoorbeeld invulvragen of combinatievragen.

Om vragen te genereren, kan de tekst worden onderverdeeld in verschillende segmenten en worden geanalyseerd om relevante informatie te identificeren. Voor meerkeuzevragen en waar/onwaar-vragen kan semantische en syntactische analyse helpen om de belangrijkste informatie te begrijpen en om van daaruit een vraag te genereren. Invulvragen en combinatievragen kunnen worden gegenereerd door delen van de tekst te selecteren en in een vraagvorm te plaatsen.

In het algemeen is het genereren van vragen met behulp van NLP-technieken een complexe taak die veel onderzoek en ontwikkeling vereist. De kwaliteit van de inputtekst is van belang voor de kwaliteit van de gegenereerde vragen. Daarom is het belangrijk om de tekst zorgvuldig te analyseren en te begrijpen voordat er vragen gegenereerd worden.

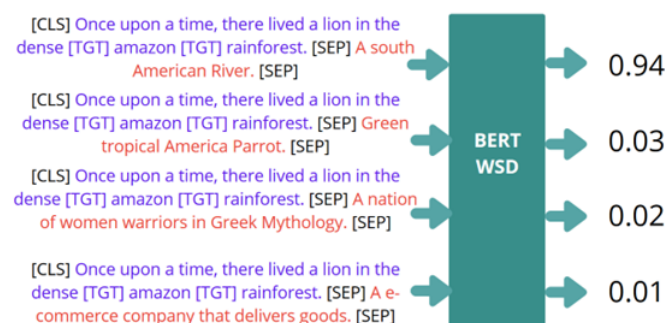
### 2.1.1. Meerkeuzevraag

Een AI-model kan worden gebruikt om een meerkeuzevraag te genereren door eerst de tekst samen te vatten met behulp van abstractive of extractive summarization. Vervolgens wordt het belangrijkste woord gebruikt om een vraag te stellen, en foutieve antwoorden te genereren met behulp van Wordnet of sense2vec. De betekenis van woorden binnen de context kan worden bepaald met het BERT WSD model. Tot slot kan een getraind T5 Transformer model vragen genereren op basis van een blok tekst en het juiste antwoord. Bij een meerkeuzevraag wordt er een lijst van mogelijke antwoorden gegeven waarbij er één of meerdere juist kunnen zijn. (AIEngineering2021)

Voor een AI-model om de belangrijkste zinnen uit de tekst te halen, wordt er gebruikt gemaakt van abstractive of extractive summarization om de tekst samen te vatten met de belangrijkste stukken.

Het AI-model moet op basis van die samenvatting een sleutelwoord vinden om er een MCQ over te stellen. Het sleutelwoord is het juiste antwoord en het model genereert een paar relevante foutieve antwoorden en de vraag.

Voor het genereren van de foutieve antwoorden gaat het model gebruikmaken van Wordnet of sense2vec in het Engels. Deze foutieve antwoorden worden ook wel *distractors* genoemd. Wordnet en sense2vec zijn libraries om te helpen gelijkaardige woorden te vinden voor het juiste antwoord. Daarnaast wordt het BERT WSD model ook gebruikt om betekenissen van woorden te achterhalen op basis van de meegegeven context. Het model geeft een zekerheidspercentage van een woord binnen de context en zijn betekenis.



Figuur 1: Het BERT WSD model

Naast foutieve antwoorden moet ook de vraag

gegenereerd kunnen worden. Er wordt hiervoor gebruikt gemaakt van een getraind T5 Transformer model dat als invoer een blok tekst krijgt waarover er een vraag gesteld wordt en het antwoord op die vraag. Bv. *“De rivier ontspringt in Peru en heeft een totale lengte van ongeveer 6.530 km. De Amazon is de tweede grootste rivier ter wereld”* en als antwoord willen we Amazon. Daarmee kan het na het trainen een vraag gaan genereren.

### 2.1.2. Waar/onwaar-vraag

Bij een waar/onwaar-vraag wordt een uitspraak gegeven waarbij er Waar of Onwaar op geantwoord moet worden. Vragen die waar zijn kunnen via abstractive of extractive summarization opgehaald worden. Het genereren van een onwaar-vraag kan op verschillende manieren:

- Veranderen van benoemde entiteit (bv. Caesar naar Nero)
- Het toevoegen of verwijderen van een negatie (bv. Het verandert ... naar Het verandert niet ... )
- Veranderen van bijvoeglijk naamwoord (bv. grootste naar kleinste)
- Veranderen van werkwoord (bv. aantrekken naar wegduwen)

Hiervoor kan Wordnet weer gebruikt worden om vragen die onwaar zijn te genereren.

### 2.1.3. Invulvraag

Een invulvraag bestaat uit een zin waarbij een woord is weggelaten en is het de bedoeling om het juiste woord hier in te vullen. Eerst worden de sleutelwoorden afgeleid uit de tekst. Dit proces wordt gedaan via de Python Keyword Extraction Library. Vanuit deze library zijn er keuzes om supervised of unsupervised modellen te gebruiken. (AIEngineering2021)

Supervised learning is een type machine learning-algoritme waarbij het model is getraind op gelabelde data, wat betekent dat de juiste uitvoer voor elk voorbeeld in de trainingdata is gegeven. Dit stelt het model in staat om voorspellingen te doen op nieuwe, ongeziene data op basis van de patronen die het heeft geleerd uit de trainingdata. In tegenstelling tot supervised learning is unsupervised learning een type machine learning-algoritme waarbij het model is getraind op ongelabelde data, wat betekent dat er geen juiste uitvoer is gegeven voor elk voorbeeld in de trainingdata. Het model moet de onderliggende structuur van de data zelf ontdekken.

Een reden om een unsupervised model te gebruiken is wanneer er geen gelabelde data beschikbaar is of wanneer het te kostbaar is om deze te verkrijgen. In deze gevallen kan unsupervised

learning worden gebruikt om patronen en structuur in de data te vinden zonder dat er gelabelde voorbeelden nodig zijn. Een andere reden om unsupervised learning te gebruiken is om het aantal gelabelde data dat nodig is voor het trainen van een supervised model te verminderen. Door unsupervised learning te gebruiken voor preprocessing van de data, kan het mogelijk zijn om een nauwkeuriger en efficiënter supervised model te trainen op een kleinere dataset. Er wordt gebruikt gemaakt van een unsupervised model omdat er geen gelabelde data beschikbaar is in dit geval.

#### 2.1.4. Combinatievraag

Bij een combinatievraag krijg je een aantal woorden en een aantal definities waarbij je ze juist moet verbinden. Voor het genereren volgt het bijna hetzelfde principe als een invulvraag. Er wordt weer gebruikt gemaakt van het Python Keyword Extraction Library om de keywords te bepalen. Een probleem hierbij is dat sommige keywords verschillende betekenissen kunnen hebben. Zoals bijvoorbeeld: Mouse kan een computermuis zijn of een dier. Dit probleem is hetzelfde als toen de meerkeuzevragen besproken en opgelost werden via het BERT WSD model als de volledige context wordt meegegeven.

### 2.2. Gamification

In een studie over de impact van gamificatie op studenten over gamificatie worden verschillende methodes besproken om gamificatie toe te passen in een educatieve omgeving. Een van de meest populaire methodes is het gebruik van een live leaderboard of top 10 scorebord. Dit kan bijvoorbeeld toegepast worden in een automatische quiz generator door de scores van gebruikers te vergelijken en ze te laten zien op een leaderboard, waardoor gebruikers gemotiveerd worden om hun prestaties te verbeteren en hoger te komen op het leaderboard. Badges, punten en rapporten om gebruikers te belonen voor het behalen van bepaalde doelen of het voltooien van specifieke taken geeft ook een positief resultaat in de betrokkenheid van de studenten. Dit kan bijvoorbeeld toegepast worden in een automatische quiz generator door gebruikers punten te geven voor het juist beantwoorden van vragen en badges te geven voor het behalen van hoge scores en op het einde van een quiz een rapport te tonen voor de gebruiker. (Smiderle2020)

Samengevat is er een verscheidenheid aan methodes die gebruikt kunnen worden in gamificatie voor een automatische quiz generator, zoals het gebruik van badges en punten, leaderboards en competitie en gegenereerde reports. Het is belangrijk om te kiezen welke methodes het best passen bij het doel van de quiz en de gebruikers om zo de meest effectieve gamificatie toe te passen.

## 3. Methodologie

Er zal een PoC gebouwd worden in de vorm van een quiz generator die de Guidance Framework kan gebruiken om vragen te genereren. Hierbij zal eerst de data in via webscraping opgehaald worden van de Guidance Framework. De PoC wordt voorgesteld via een website gemaakt met Django. Het PoC kan 4 soorten vragen (meerkeuzevragen, waar/onwaar-vragen, ...) genereren en er wordt getest of de gegenereerde vragen accuraat genoeg zijn.

Het testen van Natural Language Generation systemen is altijd een moeilijke taak geweest. Er wordt vooral gebruikt gemaakt van n-gram based similarity metrics (zoals bv. BLUE, NIST, ...) tegenwoordig wat enkel test hoeveel van een tekst voorkomt in een andere tekst. De BLUE score bleek wel een hogere score te krijgen dan de andere metrics maar er wordt nog altijd aangeraden om niet volledig hierop te steunen. (Nema2018) Er zal dus gebruikt worden gemaakt van niet alleen de BLUE score maar ook van de manuele score van een medewerker die het Guidance Framework kent. De medewerker zal een score geven van 1 tot 10 op relevantie. Dit wordt gedaan door een dataset van vragen te genereren door de PoC. De grammatica van de gegenereerde vragen en antwoorden zullen ook getest worden. Dit gebeurt met part-of-speech tagging wat een proces is waarbij elk woord in een zin geclassificeerd wordt naar zijn syntactische functie, zoals een werkwoord, naamwoord, bijvoeglijk naamwoord, etc. Dit kan gebruikt worden om de grammatica van de zin te analyseren en te verifiëren of deze correct is.

## 4. Verwacht resultaat, conclusie

Uit dit onderzoek moet blijken of een quiz generator geschikt is om vragen te genereren uit het Info Support Guidance Framework. De PoC zal een goeie indicator zijn hiervoor met een hoge BLUE score en een hoge manuele score van een medewerker van Info Support. Echter doordat er het Guidance Framework wordt gewebscraped zal een gelijkaardige structuur moeten behouden worden en in het Engels blijven geschreven worden indien er ze het Guidance Framework uitbreiden om de PoC te kunnen blijven gebruiken. Hopelijk zal de PoC kunnen gebruikt worden door Info Support om te helpen hun kennis van het Guidance Framework op te frissen.