

## [1.a]

### **Informatiebrochure voor** *Evaluating Dutch Language Models*

Geachte participant / Beste deelnemer,

U gaat deelnemen aan het onderzoek Evaluating Dutch Language Models van de Universiteit van Amsterdam faculteit Geesteswetenschappen uitgevoerd door Julia Pestel onder begeleiding van Dr. Raquel Garrido Alhama. Voordat het onderzoek begint, is het belangrijk dat u kennis neemt van de procedures die in dit onderzoek worden gevolgd. Leest u deze brochure daarom zorgvuldig door.

#### **Doel van het onderzoek**

Er zijn waarschijnlijk niet veel mensen die nog nooit van Large Language Models (LLMs) zoals ChatGPT hebben gehoord of deze hebben gebruikt. Zowel studenten als wetenschappers raadplegen LLMs om literatuur samen te vatten, te brainstormen over nieuwe ideeën en papers te schrijven. Het gebruik van deze modellen is controversieel. Een reden hiervoor is de kwaliteit van de uitvoer. Stel dat ik ChatGPT zou vragen om een artikel voor mij te schrijven, hoe zou ik dan de grammaticale correctheid van de uitvoer bepalen?

In dit onderzoek worden Nederlandse LLMs geëvalueerd op hun taalkundige kennis van verschillende grammaticale fenomenen in het Nederlands. De prestatie van de LLMs op het gebied van taalkundige kennis wordt vergeleken met die van mensen. Op deze manier kunnen we constateren in hoeverre de taalkundige kennis van LLMs overeenkomt met die van mensen. Het belang van dit onderzoek is dat het ons inzicht geeft in hoe goed LLMs grammaticale fenomenen 'begrijpen' en kunnen toepassen. Dit zal ons meer duidelijkheid geven voor welke doeleinden de verschillende LLMs geschikt zijn.

#### **Wie kan er aan dit onderzoek meedoen**

Voor dit onderzoek worden Nederlandse moedertaalsprekers uitgenodigd. Wij vragen u voorafgaand aan het onderzoek of uw eerste taal Nederlands is, of u momenteel in Nederland woont en of u taalproblemen heeft. Gezien de aard van het onderzoek is het van belang dat u een moedertaalspreker van het Nederlands bent en dagelijks gebruik maakt van de Nederlandse taal. Verder is het van belang dat u, voor zover bij u bekend, geen taalprobleem heeft, zoals dyslexie of een specifieke taalstoornis.

#### **Instructie en procedure**

Het onderzoek bestaat uit een enquête met 20 vragen. Voorafgaand aan de enquête wordt u een aantal vragen gesteld om te verifiëren dat u kan deelnemen aan het onderzoek. Deze

vragen hebben betrekking tot uw moedertaal, woonplaats en of u taalproblemen heeft. Als u voldoet aan de deelnames eisen, wordt u doorgeleid naar de enquête.

De enquête bestaat uit 20 vragen waar u bij iedere vraag twee zinnen te zien krijgt. Er wordt aan u gevraagd welk van de twee zinnen ongrammaticaal is en welke grammaticaal.

Het onderzoek zal in totaal 5 minuten duren.

### **Vrijwilligheid**

U doet vrijwillig mee aan dit onderzoek. U kunt dan ook op elk moment gedurende het onderzoek uw deelname stopzetten. Dit zal geen gevolgen voor u hebben en u bent in geen geval verplicht de hierboven beschreven procedures af te ronden. Tevens kunt u na het onderzoek altijd besluiten uw toestemming alsnog in te trekken. Als u uw deelname staakt of toestemming intrekt voordat de onderzoeksresultaten zijn gepubliceerd, worden alle tot dan toe verzamelde gegevens definitief verwijderd. Echter, als de gegevens zijn geanonimiseerd kunnen ze niet meer verwijderd worden omdat het dan niet meer mogelijk is de gegevens te herleiden naar individuele deelnemers.

### **Ongemak, risico's en verzekering**

De risico's van deelname aan dit onderzoek zijn niet groter dan die in dagelijkse situaties thuis. Uit ervaring met voorgaande, vergelijkbare onderzoeken is gebleken dat er van enig ongemak voor de deelnemers niet of nauwelijks sprake is.

Bij elk onderzoek van de Universiteit van Amsterdam geldt een standaard aansprakelijkheidsverzekering.

### **Vertrouwelijkheid van uw persoonlijke gegevens**

De gegevens die in dit onderzoek worden verzameld, zullen door de onderzoekers alleen worden gebruikt voor de doeleinden van dit project. In publicaties wordt geen gebruik gemaakt van uw persoonlijke gegevens, en u blijft onherleidbaar in alle publicaties.

De verzamelde onderzoeksgegevens zullen gecodeerd opgeslagen worden. Alleen medewerkers aan het onderzoek hebben toegang tot deze en de codering.

Er zullen geen persoonlijke gegevens worden opgeslagen of verwerkt. De niet-geanonimiseerde gegevens zullen alleen worden opgeslagen zolang het nodig is voor het onderzoek en zullen zo spoedig mogelijk worden verwijderd.

### **Rechten van betrokkenen onder de AVG**

Deelnemers kunnen op elk ogenblik bij de onderzoeker om meer informatie vragen over hun rechten als betrokkenen onder EU privacy-wetgeving, de AVG.

### **Vergoeding**

U ontvangt voor uw deelname geen vergoeding.

**Nadere inlichtingen**

Als u nog verdere informatie wilt over dit onderzoek, dan kunt u zich wenden tot Julia Pestel (e-mail: [julia.pestel@student.uva.nl](mailto:julia.pestel@student.uva.nl)) of Dr. Raquel Garrido Alhama (e-mail: [rgalhama@uva.nl](mailto:rgalhama@uva.nl); Spuistraat 134, 1012 VB Amsterdam, Nederland).

Met eventuele klachten over dit onderzoek kunt u zich wenden tot de secretaris van de Commissie Ethiek van de Faculteit Geesteswetenschappen van de Universiteit van Amsterdam; [commissie-ethiek-fgw@uva.nl](mailto:commissie-ethiek-fgw@uva.nl); Binnengasthuisstraat 9, 1012 ZA Amsterdam, Nederland.