Onderzoeksrapport

Vertaaltechnologie inzetten in chathulpverlening

Een verkennend onderzoek van CAW Halle Vilvoorde, CLB, OHUP & SAM en UAntwerpen

Esther de Boe
m.m.v. Bieke Bruijns en Luna Rodriguez,
Universiteit Antwerpen

Dit project werd mogelijk gemaakt via een Incentive voor digitale zorg, hulpverlening en ondersteuning gericht op het implementeren van een blended aanbod in het domein welzijn, volksgezondheid en gezin

Inhoudsopgave

Samenvatting 2
Inleiding 3
1. Doelstelling en onderzoeksvragen 4
2. Methodologie 4
2.1 Dataverzameling 4
2.2 Procedure livegesprekken 5
2.3 Data-analyse 5
2.3.1 Foutencategorieën 6
2.3.2 Revisie machinevertaling 6
3. Resultaten 7
3.1 Samenvatting geïdentificeerde fouten – Algemeen 7
3.1.1 Voorbeelden inhoudelijke fouten 8
3.1.2 Voorbeelden taalfouten 9
3.1.3 Vergelijking tussen de talen 10
3.2 Focus: formulering van de chattekst 12
3.3 Gebruikerservaringen 14
4. Strategieën op microniveau 15
5. Conclusies 15
5.1 Algemene bevindingen 15
5.2 Bredere tendensen op het gebied van de inzet van digitale middelen in hulpverlening/openbare diensten 16
5.3 Beperkingen van het onderzoek 16
6. Webinar en leermiddelen 17

Samenvatting

Maatschappelijke organisaties die burgers in Vlaanderen ondersteuning bieden op het vlak van mentale en fysieke gezondheid, zetten steeds vaker digitale communicatiemiddelen in. Voorbeelden hiervan zijn het Centrum voor Leerlingenbegeleiding (CLB) en het Centrum Algemeen Welzijnswerk (CAW), die gebruikmaken van chatgesprekken, anoniem of als onderdeel van blended hulpverlening. In het laatste geval gaat het om een combinatie van onsite en online afspraken. In onze huidige zeer diverse samenleving hebben deze organisaties, net zoals andere openbare instellingen, in toenemende mate te maken met taalbarrières. Om ook toegankelijk te zijn kunnen zijn voor burgers die het Nederlands (nog) niet voldoende beheersen, kwam daarom de vraag naar voren of vertaaltechnologie zou kunnen worden ingezet in de chathulpverlening. Op dit moment is namelijk nog onduidelijk in hoeverre vertaaltechnologie emotioneel geladen conversaties in realtime chathulpverlening kan ondersteunen. Bovendien wordt chattaal gekenmerkt door specifieke eigenschappen (zoals het gebruik van jargon, afkortingen, spreektaal, onjuiste spelling), waardoor de kans op vertaalfouten groot is.

In dit kader werd een eenjarig onderzoeksproject opgestart vanuit CAW Halle Vilvoorde, met als partners het CLB en de Universiteit Antwerpen, in samenwerking met Onlinehulp-Vlaanderen (OHUP), Arteveldehogeschool en het Steunpunt Mens en Samenleving (SAM). Het project werd mogelijk gemaakt via een incentive voor blended hulpverlening van het Departement Zorg van de Vlaamse Overheid. In dit onderzoeksrapport wordt verslag uitgebracht van de resultaten van dit verkennende project, waarin de kwaliteit van de vertaaltechnologie DeepL van authentieke chatsessies werd onderzocht. Op basis van een aangepaste versie van een raamwerk voor kwaliteitsmeting van vertalingen (TAUS Dynamic Quality Framework/Multidimensional Quality Metrics) werd de kwaliteit van machinevertalingen in het Frans, Arabisch en Turks van 20 bestaande anonieme chattranscripties en 8 realtime chatsessies geëvalueerd. Ook werd de tevredenheid van deelnemers aan de realtime chatsessies onderzocht.

De analyses van de machinevertalingen van de chattranscripten laten zien dat de kwaliteit van de gebruikte vertaaltechnologie nog veel te wensen over laat en varieert tussen de verschillende talencombinaties. Het grootste deel van de vertaalfouten bleek gerelateerd aan de inhoud en vorm van de brontekst. Zo zorgde het gebruik van gespecialiseerde terminologie en afkortingen in de context van lokale openbare diensten zoals onderwijstermen, alsook het gebruik van een informeel, figuurlijk taalregister, onduidelijke formuleringen en korte chatregels met grammaticaal onvolledige zinnen voor onjuiste machinevertalingen. Dit resultaat impliceert echter ook dat een aantal frequente vertaalproblemen vermeden kan worden indien meer aandacht wordt besteed aan correcte formuleringen en het toevoegen van gespecialiseerde terminologielijsten in de vertaaltechnologiesoftware. Wel moet rekening worden gehouden met het feit dat deze verbeteringen vooral aan de kant van de hulpverleners zullen plaatsvinden en minder aan de kant van de cliënten. Uit de analyses van de livegesprekken bleek namelijk dat hulpverleners het communicatieproces nauwlettend in de gaten houden en via metacommunicatie het gesprek bijsturen wanneer ze twijfels hebben over de accuraatheid van de vertaling.

Wat het gebruik van de vertaaltechnologie betreft, waren de medewerkers in het algemeen tevreden over het gebruik van DeepL in hun chatsessies. Ondanks dat de werkwijze (kopiëren en plakken tussen het chatvenster en het DeepL-venster) als redelijk omslachtig en vertragend werd beschouwd, zien medewerkers het gebruik van vertaaltechnologie als een meerwaarde om snel taalbemiddeling in te kunnen zetten en direct contact te hebben met hun cliënten. Wel wijzen zij erop dat het gebruik van vertaaltechnologie eerder bruikbaar is voor gesprekken van praktische aard en met gekende cliënten dan voor emotioneel beladen gesprekken en contacten met anonieme cliënten.

Deze resultaten vormen een belangrijk uitgangspunt bij de ontwikkeling van een handelingskader voor het inzetten van vertaaltechnologie in de chathulpverlening en de besluitvorming rond de integratie hiervan in de chatsoftware die gebruikt wordt door CAW en CLB. Ook werden op basis van deze resultaten een webinar georganiseerd en leermiddelen ontwikkeld in verband met het gebruik van vertaaltechnologie in chathulpverlening voor medewerkers van de genoemde organisaties en andere organisaties voor openbare dienstverlening.

Inleiding

In het voorliggende rapport wordt verslag gedaan van het project "Inzetten van vertaaltechnologie in chathulpverlening". Dit project is tot stand gekomen op vraag van de partijen CLB (Centrum voor Leerlingenbegeleiding)/CAW (Centrum Algemeen Welzijnswerk), met steun van SAM (Steunpunt Mens en Samenleving) en Onlinehulp-Vlaanderen (Arteveldehogeschool), die instaan voor de ondersteuning en ontwikkeling van (chat)hulpverlening in Vlaanderen. Zij hebben de Universiteit Antwerpen (Faculteit Letteren & Wijsbegeerte - Departement Toegepaste Taalkunde/Vertaal- en Tolkwetenschap) gecontacteerd om samen een manier te vinden om te onderzoeken of vertaaltechnologie in de vorm van machinevertaling (in de Vertaalwetenschap "machine translation" of "MT" genoemd) kan worden ingezet in het specifieke domein van (blended) chathulpverlening. Doordat de chathulpverlening op dit moment uitsluitend in het Nederlands wordt aangeboden, hebben cliënten die het Nederlands niet of slechts in beperkte mate machtig zijn, nauwelijks of geen toegang tot deze dienstverlening. Daarom kwam de vraag naar voren of het mogelijk is MT in te zetten in chatgesprekken, zodat deze toch aanzienlijke groep cliënten ook gebruik kan maken van online hulpverlening. De belangrijkste factoren bij het beoordelen van de geschiktheid van MT zijn de kwaliteit van de vertalingen en de tevredenheid onder de gebruikers.

Na inventarisatie van de specifieke noden aan wetenschappelijke kennis rond het combineren van chatgesprekken met vertaaltechnologie, werd in overleg besloten een aanvraag te doen voor een incentive van de Vlaamse overheid - Departement Zorg, voor een eenjarig project binnen het kader "Incentives voor digitale zorg, hulpverlening en ondersteuning, gericht op het implementeren van een blended aanbod in het domein welzijn, volksgezondheid en gezin"[1]. Deze aanvraag werd formeel ingediend door CAW Halle Vilvoorde in juni 2024, met als partners CLB en UAntwerpen. Begin oktober 2025 werd de subsidie toegekend als een van vijf gefinancierde projecten binnen deze incentive. Vanaf oktober begon de dataverzameling in 2 fases, waarbij in fase 1 door CAW en CLB bestaande anonieme transcripten werden aangeleverd en in fase 2 de transcripten van de live chatgesprekken. Voordat van start werd gegaan met de dataverzameling werd toestemming gevraagd en verkregen via de Ethische Commissie Sociale en Humane Wetenschappen van de UAntwerpen[2]. Ook werd er een datamanagementplan opgesteld en goedgekeurd (zie bijlage 1). De deelnemers aan de livegesprekken werden op voorhand geïnformeerd over de onderzoeksprocedure, waarna hun toestemming om deel te nemen werd verkregen via een standaardformulier. Deze documenten worden bewaard door de organisaties CAW en CLB. Daarnaast hebben de twee studenten in de Master Tolken die in het kader van hun masterscriptie hebben bijgedragen aan het ontwikkelen van de methodologie en de data-analyse, een geheimhoudingsovereenkomst ondertekend om de dataveiligheid te garanderen.

Leeswijzer

In dit onderzoeksrapport worden eerst de doelstelling en de onderzoeksvragen besproken die centraal stonden in het project (paragraaf 1). Vervolgens wordt toegelicht hoe deze vragen onderzocht worden door middel van een toelichting bij de methodologie, dat wil zeggen, de dataverzameling, procedure en data-analyse (paragraaf 2). Vervolgens worden de belangrijkste resultaten weergegeven door middel van een samenvatting van de geïdentificeerde fouten, met een algemeen deel en een deel waarin de focus ligt op een specifieke problematiek die werd geobserveerd, en een deel met de gebruikerservaringen (paragraaf 3). Daarop volgt een voorstelling

[1] https://www.departementzorg.be/nl/digitale-zorg-hulpverlening-en-ondersteuning
[2] Goedkeuring Commissie Ethische advisering Sociale en Humane Wetenschappen Universiteit Antwerpen, dossier SHW_2024_204_2

van mogelijke strategieën op microniveau voor een zo verantwoord mogelijk gebruik van vertaaltechnologie in de chathulpverlening (paragraaf 4). Het rapport wordt afgesloten met conclusies van het huidige onderzoek, inclusief bredere tendensen en beperkingen van dit onderzoek en suggesties voor vervolgonderzoek (paragraaf 5), en een overzicht van de concrete output van het project in de vorm van het webinar dat plaatsvond op 15 september 2025 en de ontwikkelde leer- en communicatiemiddelen (paragraaf 6).

1. Doelstelling en onderzoeksvragen

Het doel van het project was om inzicht te krijgen of en in welke mate MT op een verantwoorde manier kan worden ingezet in (blended) chathulpverlening, waarbij de resultaten van het onderzoek direct aansluiten vinden in de praktijk. Dit gebeurde in de vorm van het ontwikkelen van leermiddelen en de organisatie van een webinar. Hiermee worden medewerkers op een hands-on manier ondersteund bij het gebruik van MT in chathulpverlening met anderstaligen.

Bij dit onderzoek gingen we uit van de volgende hoofdvraag:

- Wat is de kwaliteit van machinevertaling via DeepL van bestaande anonieme chatgesprekken en van live chatgesprekken Nederlands-Arabisch/Frans/Turks in het kader van online hulpverlening binnen de organisaties CLB en CAW?

De deelvragen luidden als volgt:

- Wat zijn veel voorkomende fouten in de machinevertaling van de transcripten en zijn er oorzaken van deze fouten aan te wijzen?
- Kunnen er oplossingen voor bepaalde fouten worden gevonden en zo ja, welke?

2. Methodologie

2.1 Dataverzameling

De onderzoeksvraag werd beantwoord op basis van kwalitatieve analyses van transcripties van DeepL-Translator-machinevertalingen (hierna: “DeepL”) tussen het Nederlands en het Frans, Arabisch en Turks. DeepL[3] is een tool voor vertaaltechnologie die werkt op basis van zogenaamde “neural machine translation”, wat wil zeggen dat de tool gebruikmaakt van een neuraal netwerk van artificiële intelligentie dat (ver)taalpatronen herkent op basis van grotere (parallelle) tekstcorpora, om zo nieuwe vertalingen te genereren. De geanalyseerde data bestond uit fragmenten van bestaande (Nederlandstalige) anonieme chats (“statische transcripten”) en van live chatgesprekken Nederlands-Arabisch/Frans/Turks (“dynamische transcripten”) die werden gehouden in het kader van blended hulpverlening met gekende cliënten. Deze transcripten werden geanonimiseerd aangeleverd door de organisaties CAW en CLB via de beveiligde dataverzending Belnet[4]. In totaal werden er 20 statische transcripten (dataset 1) en 8 dynamische transcripten (dataset 2) aangeleverd. De transcripten in dataset 1 werden door een onderzoeksmedewerker (een

[3] https://www.deepl.com/en/products/translator
[4] https://www.belnet.be/nl

masterstudent Tolken) van de UAntwerpen vertaald via DeepL, regel voor regel, zoals dit in een live chatgesprek ook zou gebeuren. Hiervoor werd een betaald account gecreëerd om dataveiligheid te maximaliseren. Dat geldt ook voor de DeepL-accounts die de medewerkers van CAW en CLB gebruikten in hun experimentele gesprekken. In totaal werden hiervoor 7 betaalde accounts aangemaakt. Daarnaast werd een Qualtrics-enquête opgesteld met vragen over de livegesprekken (zie bijlage 2). Deze werden echter door slechts twee medewerkers ingevuld. Sommigen gaven liever een toelichting via e-mail, anderen via een medewerker van SAM. Deze ervaringen (dataset 3) zijn gebundeld en worden bij de resultaten weergegeven.

2.2 Procedure livegesprekken

Het softwaresysteem dat door de organisaties CAW en CLB voor de chathulpverlening gebruikt wordt is de zogenaamde Social Intervention Tool oftewel "SIT-tool" (SIT – Secure Chat Software for Social Orgs, eHealth and Government[5]). Het is momenteel om dataveiligheidsredenen niet mogelijk om via API een extensie voor MT direct te integreren in de SIT-tool. Daarom gebruikten de medewerkers die deelnamen aan de experimentele gesprekken Deep L door middel van het kopiëren naar DeepL van hun chatteksten, en vervolgens het plakken van de vertaling in de chattool (zie bijlage 3). Ook ontvingen de medewerkers instructies in de vorm van een document en een instructievideo over hoe zij de transcripten na afloop van het gesprek konden downloaden vanuit de SIT-tool en hoe zij deze moesten anonimiseren.

2.3 Data-analyse

Alle transcripten werden geannoteerd met behulp van een schema voor kwaliteitsmeting dat is opgesteld op basis van het TAUS Dynamic Quality Framework (DQF)/Multidimensional Quality Metrics (MQM)[6]. Door middel van de identificatie van fouten in verschillende subcategorieën en hun impactniveau (ernst van de fouten) laat dit schema zien hoe adequaat de MT is qua inhoud en taal.

		Insignificant	Minor	Moderate	Major	Critical	**TOTAL**
Accuracy	Addition						
	Omission						
	Mistranslation						
	Over-translation						
	Under-translation						
	Untranslated						
	TOTAL						
Fluency	Punctuation						
	Capitals						
	Spelling						
	Grammar						
	Gramm. register						
	Inconsistency						
	Unidiomatic						
	TOTAL						
Terminology	Specific						
	Inconsistency						
	TOTAL						
	TOTAL						

Tabel 1. *Schema voor kwaliteitsmeting op basis van TAUS DQF/MQM*

[5] https://sittool.com/ (Voor een demo zie: https://sam.sittool.net/)
[6] https://themqm.org/

2.3.1 Foutencategorieën

De hoofdcategorie 'Accuracy', brengt de accuraatheid van de machinevertalingen in kaart en is onderverdeeld in zes subcategorieën. De categorie 'Addition' is voorzien voor vertaalfouten waarbij de doeltekst een woord(groep) of zin bevat die niet aanwezig was in de brontekst. Hiertegenover staat de categorie 'Omission', wat inhoudt dat er een woord(groep) of zin die wel aanwezig was in de brontekst ontbreekt in de MT. Een fout wordt in de categorie 'Mistranslation' geplaatst wanneer de doeltekst geen accurate vertaling voorziet van een element in de brontekst. Daarnaast geven de categorieën 'Over-translation' en 'Under-translation' respectievelijk aan of de doeltekst specifieker of minder specifiek is dan de brontekst. Wanneer een element van de brontekst letterlijk overgenomen wordt in de MT zonder vertaald te zijn, valt dit in de categorie 'Untranslated text'.

De tweede hoofdcategorie is 'Fluency'. Binnen deze categorie onderscheiden we zeven subcategorieën. Een daarvan is 'Punctuation', waaronder toevoegingen, weglatingen of aanpassing van leestekens vallen. Daarnaast werden spelfouten in de MT ondergebracht in de subcategorie 'Spelling'. Grammaticafouten zoals onjuiste werkwoordtijden, zinsbouw en foute grammaticale geslachten vallen onder 'Grammar', terwijl registerfouten (bijv. "u" in de MT terwijl er in de brontekst "jij" staat) onder 'Grammatical register'. Tot de subcategorie 'Inconsistency' behoren fouten die de interne consistentie binnen de tekst verstoren (een en de dezelfde term wordt op verschillende manieren vertaald). Aanpassingen in hoofdlettergebruik behoren tot de subcategorie 'Capitals'. Tot slot bevat de subcategorie 'Unidiomatic' tekstelementen die geen idiomatisch equivalent krijgen in de doeltaal.

De derde hoofcategorie is 'Terminology', waarbinnen onjuist vertaalde specifieke termen (bijv. binnen de CLB-/schoolcontext) vallen in de subcategorie 'Specific terminology'. Wanneer specifieke terminologie op inconsistente wijze wordt vertaald is dit 'Inconsistency'. Indien dat bepaalde begrip verder in dezelfde tekst op een andere manier vertaald wordt, wordt dit niet meer als fout tegen specifieke terminologie meegerekend.

De impactniveaus variëren van 'insignificant' – 'minor' – 'moderate' – 'major' – 'critical', waarbij 'insignificant' aangeeft dat verwacht wordt dat de geannoteerde fout geen impact heeft op het begrip van de tekst en nauwelijks impact bij minor, terwijl bij moderate er een lichte impact en bij major een grote impact verwacht kan worden op het tekstbegrip. 'Critical' wil zeggen dat de vertaalfout het tekstbegrip volledig in de weg staat en tot ernstige miscommunicatie kan leiden.

2.3.2 Revisie machinevertaling

Terwijl de Franse vertalingen door de UAntwerpen intern konden worden gereviseerd, waren we voor de Turkse en Arabische revisies van de machinevertalingen aangewezen op verschillende externe vertalers, ingeschakeld via het Sociaal Vertaalbureau. Ondanks de door ons opgestelde richtlijnen voor de revisie (zie bijlage 4), bleek het toch onmogelijk om alle transcripten op identieke wijze en op een even gedetailleerd niveau te laten reviseren. Door dit verschil in gradatie (gedetailleerde revisie versie van het Frans en een eerder ruime beoordeling van de fouten en suggesties ter verbetering in het Turks en Arabisch) was het niet mogelijk om een kwantitatieve vergelijking te maken in het aantal en type fouten tussen de talen. Daarom hebben we ervoor gekozen om op basis van de Franse revisies de belangrijkste fouten te identificeren en deze te vergelijken met dezelfde fragmenten in het Arabisch en het Turks. Dit werd ook gedaan voor het Turks en het Arabisch, om zo vast te kunnen stellen welke problemen in alle talen voorkomen en welke taalspecifiek zijn.

Daarnaast werd er in de analyses rekening gehouden met de aard van de twee verschillende datatypes, dat wil zeggen, enerzijds de bestaande, statische transcripties en anderzijds de live gesprekken. In het eerste geval werd vooral gekeken naar het soort fouten en de oorzaak hiervan, terwijl in de het tweede geval ook kon worden onderzocht welke impact eventuele vertaalfouten hadden op de communicatie door te bekijken of deze werden gesignaleerd door de communicatiepartners en, zo ja, hoe hiermee hier vervolgens werd omgegaan.

3. Resultaten

Hieronder worden samenvattingen gegeven van frequente fouten per hoofdcategorie, met een speciale focus op fouten die gerelateerd zijn aan de formulering van de chattekst, inclusief illustratieve voorbeelden uit de verschillende talen. In de voorbeelden worden steeds de originele chattekst weergegeven, de machinevertaling via DeepL en de letterlijke vertaling van de machinevertaling door de revisor. Onder elke tabel staat uit welk chatgesprek het voorbeeld afkomstig is en het impactniveau van de fout.

3.1 Samenvatting geïdentificeerde fouten – Algemeen

- De analyses van de vertaling van de anonieme chattranscripten in Turks, Arabisch, Frans en (deels) Spaans laten zien dat er op dit moment **veel problemen zijn op het gebied van inhoudelijke correctheid (*accuracy*) en taal (*fluency*)**. Er is geen enkel vertaald transcript zonder **fouten**.
- Van veel van de **inhoudelijk en taalkundig** onjuiste vertalingen is de betekenis wel **min of meer begrijpelijk uit de context** van het gesprek. Dit heeft te maken met het **pragmatische aspect** van taal: veel taalgebruikers zijn goed in staat om een boodschap te begrijpen, ondanks incomplete zinnen en typfouten in de brontekst. Een menselijke vertaler zou deze fouten corrigeren in de vertaling, maar tot deze typisch **menselijke optimalisatie** van incomplete of incorrecte brontekst is MT slechts in beperkte mate in staat. Ook is het van belang dat er niet per definitie van worden uitgegaan dat **hulpvragers** in staat zijn **de inhoud van de vertaalde boodschap juist te interpreteren**, vooral gezien de **kwetsbare situatie** waarin sommigen van hen zich bevinden.
- Fouten worden zeer frequent veroorzaakt door een **gebrek aan context, onjuistheden in de spelling of uitdrukking in de brontaal** (bijvoorbeeld spelfouten, dialect, spreektalige elementen of weglatingen van het grammaticale onderwerp), **afkortingen** en **vakjargon**.
- In dit verband moet ook rekening worden gehouden met het feit dat de anderstaligen die deelnemen aan de chatgesprekken in veel gevallen **meertalig** zijn en **de geletterdheid in de** dominante taal of de thuistaal **vaak spreektalig is,** wat kan leiden tot fouten op het gebied van spelling, interpunctie, grammatica en idioom in de chatteksten die ze produceren, met ook weer een **verhoogd risico op een onjuiste machinevertaling**.
- Fouten kunnen ook gerelateerd zijn aan **timing:** soms is het niet duidelijk of een chatregel een antwoord is op een vorige chatregel, als de medewerker en de cliënt tegelijkertijd een boodschap intypen. In verband met het goed kunnen volgen van de logica, moeten beide communicatiepartners erop letten pas een bericht in te typen als de ander niet meer aan het typen is (dit is zichtbaar in de software).
- Daarnaast werd in de livegesprekken geobserveerd dat wanneer er een **opvallend vreemde MT** wordt geproduceerd, **cliënten die vrijwel steeds lijken te negeren** (er komt geen reactie op of vraag over) terwijl wanneer **hulpverleners** een dergelijke vertaling tegenkomen, zij meestal **wel reageren** door een **bijvraag te stellen** om op die manier de **juiste betekenis van de boodschap**

van de cliënt te kunnen achterhalen. Uit de interviews met de medewerkers na afloop bleek wel dat de **deelnemende cliënten vaak het Nederlands (deels) machtig zijn** en zij de beide talen in chatregels zagen, wat deze observatie ook zou kunnen verklaren.

3.1.1 Voorbeelden inhoudelijke fouten

In deze hoofdcategorie zien we met name de subcategorieën *mistranslation* (inhoudelijk onjuiste vertalingen) en *omission* (weglatingen). Bij *mistranslation* zien we vaak fouten met een groot impactniveau, terwijl het bij *omission* meestal gaat om fouten met weinig impact op de begrijpelijkheid en overbrengen van de inhoud van de brontekst.

Een voorbeeld van *mistranslation*:

Chattekst	Machinevertaling
14:07 Operator: Aha, oké, maar ik hoor dat je een jongen hebt die kut doet tegen jou... Das rot... Vertel me daar eens wat meer over?	آها، حسناً، ولكنني سمعت أن لديك ولداً يمارس الجنس معك.... Das rotten... أخبرني المزيد عن ذلك؟ **Letterlijke vertaling** Aha, oké, maar ik hoor dat je een jongen hebt die heeft sex met jou vertel me daar eens wat meer over? … Das rotten

CAW-10-R 5, Arabisch – Impactniveau critical

In het bovenstaande voorbeeld zijn we een ernstige *mistranslation* van de uitdrukking 'kut doen'. De betekenis is totaal anders dan bedoeld en kan voor ernstige misverstanden zorgen. Behalve deze inhoudelijke fout zien we ook taalkundige fouten. De uitdrukking [sic] 'Das rot' is onvertaald overgenomen in de MT. Door de aanwezigheid van Latijnse schrift wordt de zinsvolgorde in de MT aangepast, een veelvoorkomende fout bij vertalingen uit het Arabisch die we vooral constateren wanneer er een afkorting in Latijns schrift (bijv. CLB) wordt overgenomen.

Zoals vermeld gaat het bij omissie veelal om fouten met een laag impactniveau. Toch zijn er ook een paar voorbeelden waarbij essentiële informatie wegvalt. Een voorbeeld van een dergelijke **omissie** wordt hieronder gegeven:

Chattekst	Machinevertaling
18:12 Operator: De clb-contactpersoon van de school kan daar meer over vertellen.	La personne de contact de l'école peut vous renseigner à ce sujet. **Letterlijke vertaling** De contactpersoon van […] de school kan u meer informatie geven over dat onderwerp.

CLB-2-R28, Frans – Impactniveau major

We zien in dit voorbeeld dat 'clb' is weggelaten in de MT, waardoor het onduidelijk wordt voor de cliënt dat het iemand is van CLB die gecontacteerd moet worden, en niet zomaar een contactpersoon van de school.

De overige categorieën (addition, under-translation, over-translation en untranslated) kwamen relatief weinig voor en hadden over het algemeen een laag impactniveau.

3.1.2 Voorbeelden taalfouten

Binnen de categorie Taalproblemen (*Fluency*) zijn er veel **grammaticale issues:**

- **Inconsistentie geslacht**: In het Arabisch en Frans, waarin het geslacht van taalgebruiker wordt uitgedrukt in de verbuiging van adjectieven of werkwoordsvormen, wordt dit inconsistent toegepast. Eerst wordt bijvoorbeeld de vrouwelijke uitgang gebruikt ("vous êtes **contente**") en een paar regels verder de mannelijke ("on vous a **prévenu** ?")
- **Bias:** in een Frans-Nederlandse livegesprek (CLB_06-R6) wordt de genderneutrale term "verpleegkundige" vertaald met "infirmière" (de vrouwelijke vorm dus) terwijl het gender van de medewerker nog niet bekend is uit de context.
- **Grammaticaal register en inconsistente toepassing:** in de meeste gevallen (met name in het Frans en het Turks) worden aansprekingen met "je" in de brontekst aangepast naar de u-vorm in de MT. Dit is niet echt een probleem omdat bijvoorbeeld in het Frans in professionele relaties de u-vorm meer gebruikelijk is dan de jij-vorm. Omgekeerd komt het ook voor dat u-aansprekingen in de Franse brontekst in de Nederlandse MT worden omgezet naar jij-vormen. Wel verwarrend en mogelijke negatief voor de therapeutische relatie tussen de hulpverlener en de hulpvrager is dat de verschillende aanspreekvormen (jij- en u-vorm) regelmatig door elkaar heen worden gebruikt. In onderstaan voorbeeld hiervan zien we overigens ook een idiomatische fout ("travailles" i.p.v. "sur lesquelles tu travailles") en een ontbrekend leesteken (twee uitroeptekens in de brontekst, 1 in de MT):

Chattekst	Machinevertaling
19:49 Operator: Maar wel goed dat **je** openstaat voor hulp!! En dit zijn ook de zaken waar **je** met **je** psychotherapeut rond werkt?	Mais c'est une bonne chose que **vous soyez** ouvert à l'idée d'obtenir de l'aide ! Et ce sont aussi les choses que ***tu travailles*** avec **ton** psychothérapeute ?
	Letterlijke vertaling
	Maar wel goed dat **u** openstaat voor hulp! En dit zijn ook de zaken waar **je** met **je** psychotherapeut [rond] werkt?

CLB4-R5, Frans – Impactniveau moderate

- **Veelvoorkomende afkortingen** (zoals CLB!) zorgen voor fouten en inconsistenties in de MT. Met name in het Arabisch leiden letterwoorden of afkortingen in het Nederlands in de MT tot onjuiste aanpassingen in de zinsvolvolgorde in de machinevertaling. Bij alle talen stellen we vast dat er voor de specifieke afkortingen en acroniemen (en ook voor andere termen uit het schooljargon, socialezekerheidsstelsel en medische termen) vaak verschillende vertalingen worden geproduceerd voor dezelfde afkorting of term, dus inconsistentie in de machinevertaling:

Chattekst	Machinevertaling
18:13 Operator: … Dan kan men je mail naar jouw **CLB-begeleider** doorsturen.	… Il pourra alors transmettre votre courriel à votre **superviseur de NCLC**.
	Letterlijke vertaling
	… Dan zal men uw mail naar de toezichthouder van het NCLC doorsturen.

CLB2-R32, Frans – Impactniveau: major

- **Specifieke terminologie/culturele referenties:** begrippen die typisch zijn voor de Vlaamse context/cultuur worden onjuist of letterlijk vertaald waardoor de betekenis niet duidelijk is. Hieronder vallen ook begrippen die typisch zijn voor de Vlaamse schoolcontext, zoals "een naamloos jaar". Voorbeeld:

Chattekst	Machinevertaling
18:12 Operator: Vb. **zorgboer** of **time-out** buiten de school?	Par exemple, **l'agriculteur qui s'occupe des enfants** ou **le temps mort** en dehors de l'école ?
	Letterlijke vertaling
	Bijvoorbeeld, een **boer die zich bezighoudt met kinderen** of een pauze buiten de school?

CLB2-R27, Frans – Impactniveau: major

- Daarnaast komen er ook volledig **arbitraire fouten** voor, waarvan de oorzaak niet meteen duidelijk is, zoals in het voorbeeld hieronder. De zeer gangbare uitdrukking 'klinken als' wordt te letterlijk vertaald, wat de output van de MT moeilijker te begrijpen maakt dan de brontekst:

Chattekst	Machinevertaling
15:07 Hulpverlener: Oei <u>dat klinkt</u> inderdaad als een moeilijke situatie	**Kulağa gerçekten de** zor bir durum gibi geliyor.
	Letterlijke vertaling
	Het komt naar het oor als echt een moeilijke situatie.

CAW6-R2, Turks – Impactniveau: moderate

3.1.3 Vergelijking tussen de talen

Binnen de statische transcripten keken we naar fouten die werden geannoteerd als 'major' en 'critical'. Om zo dicht mogelijk te blijven bij een werkelijk scenario, lag de focus hierbij op de vertalingen van de chatregels die door de hulpverlener werden ingetypt, omdat die in de reële situatie vertaald zouden worden naar de vreemde taal. De foutieve machinevertalingen die we hierin vonden, vergeleken we met de machinevertalingen in de andere talen. Hieronder volgt een selectie

van voorbeelden. De letterlijke vertalingen zijn gemarkeerd in rood wanneer ze onjuist zijn en in het groen wanneer ze wel juist zijn.

In een eerste voorbeeld (CAW1-R23) zien we een *major* fout in zowel het Frans als het Arabisch na het gebruik van een acroniem in de brontekst. In de Turkse machinevertaling is het acroniem wel juist overgenomen. Het fragment betreft een term die naar verwachting regelmatig voorkomt in de chathulpverlening via CLB en CAW en kan worden vermeden door te werken met glossaria waarin veelvoorkomende termen met hun vertaling kunnen worden opgenomen. Binnen DeepL en andere betalende vertaaltools is het mogelijk om dergelijke glossaria op te nemen.

14:12 Hulpverlener: Zou ze eventueel openstaan om deze zomer eens te gaan babbelen met iemand van het JAC? Zodat ze een beetje kan vertellen wat er in haar omgaat?	
Machinevertaling	**Letterlijke vertaling**
Serait-elle prête à discuter avec un membre du CCM cet été ? Pour qu'elle puisse partager un peu de ce qui la préoccupe ?	Zou ze … openstaan om … te praten met iemand van het **CCM** deze zomer? Zodat ze een beetje kan delen over wat er in haar omgaat?
هل من الممكن أن تكون منفتحة على إجراء محادثة مع شخص ما من مركز اليابان للمساعدة القانونية هذا الصيف؟ حتى تتمكن من مشاركة بعض ما يدور في ذهنها؟	Zou ze openstaan voor een gesprek met iemand van het **Japan Legal Aid Centre** deze zomer? Zodat ze wat van haar gedachten kan delen?
Bu yaz JAC'den biriyle sohbet etmeye açık olabilir mi? Böylece aklından geçenleri biraz paylaşabilir mi?	Zou ze deze zomer openstaan voor een gesprek met iemand van **JAC**? Zodat ze haar gedachten eens kan delen?

CAW1-R23

Een tweede voorbeeld (CAW-5-R23) betreft een major mistranslation. Ook hier zien we verschillen tussen bepaalde talen en overeenkomsten tussen andere. In het Frans wordt de chatregel juist vertaald, maar in het Arabisch en Turks wordt een vergelijkbare, foutieve vertaling gegeven door DeepL. Dit is een voorbeeld van een fout die ernstig is wanneer we die geïsoleerd van de rest van het gesprek bekijken, maar waarvan de betekenis in de context van het gesprek eventueel toch te begrijpen is. Zoals eerder vermeld is het echter steeds belangrijk om er niet automatisch vanuit te gaan dat cliënten op basis van context toch in staat zijn een foutieve vertaling te begrijpen. In alle talen zien we daarnaast het gebruik van de u-vorm in plaats van de jij-vorm.

17:08 Hulpverlener: en is er nog iets afgesproken qua vervolg opvolging na je **opname**?	
Machinevertaling	**Letterlijke vertaling**
D'autres mesures ont-elles été convenues en termes de suivi après votre admission ?	Zijn er nog andere maatregelen overeengekomen over de opvolging na uw **opname**?
وهل تم الاتفاق على أي شيء آخر من حيث المتابعة بعد قبولك؟	Werd er nog iets anders afgesproken in termen van follow-up na uw **aanvaarding**?
Ve kabulünüzden sonra takip açısından başka bir şey kararlaştırıldı mı?	En is er nog iets afgesproken in termen van follow-up na uw **aanvaarding**?

CAW-5-R23

Tot slot nog een voorbeeld waarbij het in alle drie de talen fout gaat. De uitdrukking die door de hulpverlener wordt gebruikt, wordt niet correct opgepikt door DeepL. In zowel het Frans, Arabisch als Turks wordt de uitdrukking letterlijk vertaald, waardoor de boodschap zo goed als onbegrijpelijk wordt. Hier is het dus zaak voor de hulpverlener om eenvoudiger en minder figuurlijk taalgebruik te bezigen. Ook hier zien we de eerdergenoemde en veel voorkomende fout voorkomen dat de u-vorm wordt gebruikt in plaats van de jij-vorm (Frans en Turks).

14:06 Hulpverlener: Als je zegt dat je **op bent,** wat bedoel je daar dan mee?	
Machinevertaling	**Letterlijke vertaling**
Quand vous dites que vous êtes debout, qu'entendez-vous par là ?	Als u zegt dat u **rechtop** bent, wat bedoelt u daar ... mee?
عندما تقولين أنك فوق، ماذا تقصدين بذلك؟	Als je zegt dat je **aan de top staat**, wat bedoel je daar dan mee?
Yukarıdayım derken ne demek istediniz?	Wat bedoelt u met **ik ben boven**?

CAW-4-R16

3.2 Focus: formulering van de chattekst

Zoals hierboven vermeld, heeft de manier waarop de chatteksten worden geformuleerd een grote invloed op de kwaliteit van de vertalingen. Omdat dit iets is waar de deelnemers in de chatgesprekken (met name de hulpverleners) zelf actief mee om kunnen gaan, focussen we ons op deze categorie, zodat op basis hiervan praktische strategieën kunnen worden opgesteld die deelnemers kunnen helpen bij het optimaliseren van hun chatcommunicatie. Hieronder een opsomming van de problemen op dat gebied:

- **Dubbelzinnige** formuleringen kunnen leiden tot fouten in de MT, bijv. "[…] omdat ze dit niet goed zien komen" leidt tot een onbegrijpelijke letterlijke vertaling in het Arabisch (CAW1- AR, 15) of **specifieke uitdrukkingen** zoals "Kropt ze veel op?" (wordt letterlijk vertaald in het Arabisch als "Blaast ze veel dan?" (CAW1-AR, 16).
- **Ongebruikelijke samenstellingen**

Chattekst	**Machinevertaling**
14:41 OPERATOR: nooit bij dans-stress bij een voorstelling of zo?	لا ترقص أبداً في الرقص-التوتر في الأداء أو شيء من هذا القبيل؟
	Letterlijke vertaling
	Dans nooit in een dans stress bij een voorstelling of zoiets?

CLB5-R45, Arabisch – Impactniveau moderate

- **Figuurlijk taalgebruik of uitdrukkingen die minder gangbaar** zijn leveren problemen op in die vertaling, waarbij de MT twee woorden afzonderlijk vertaalt en die dan samenbrengt, terwijl die combinatie in de doeltaal niet bestaat of voor een andere betekenis zorgt (bijv. 'het slachtoffer uithangen' →' traîner la victime' = (letterl.) 'het slachtoffer voortslepen (CLB6-R31). Zie ook voorbeeld (livegesprek) CLB-03 (Arabisch): "Het is voor XXX de innerlijke onrust die hem parten speelt op school". Dit leidt tot de onjuiste vertaling "Het is voor XXXX deze geïnternaliseerde stoornis die hij soms omzeilt op school".

- **Dialectwoorden**, hoe gebruikelijk deze ook zijn in de dagelijkse spreektaal, komen er in de MT vaak in een onjuiste betekenis uit. Termen als “op kot gaan” en “seffes” krijgen een **totaal andere betekenis in de MT** of worden **letterlijk overgenomen,** waardoor de boodschap **onbegrijpelijk** wordt. Zo wordt ‘op kot gaan’ in het Arabisch vertaald met ‘thuis’, en in het Frans met ‘opgravingen’.
- **Structuur van de chattekst en timing:**

- De MT wordt verstoord wanneer **doorlopende informatie over meerdere chatregels wordt verdeeld** en de losse regels niet meer grammaticaal zijn, met name wanneer in de vervolg-chatregel het onderwerp ontbreekt. Dit wordt geïllustreerd in het voorbeeld hieronder. In regel 23 eindigt de hupverlener met een zin, die door de hulpvrager wordt opgevat als een volledige zin, omdat de chatregel wordt ingegeven en ‘meisje’ direct aansluitend hierop antwoordt. Vervolgens blijkt in regel 25 dat de boodschap van de hulpverlener nog niet compleet was, maar verdeeld over 2 chatberichten. In regel 25 voegt de hulpverlener een betrekkelijke bijzin toe die op het laatste deel van regel 23 slaat, namelijk ‘je psycholoog’. De machinevertaling, die deze koppeling niet kan maken, vertaalt deze toch belangrijke informatie niet (*omission*).

Nr.	Chatregel	Machinevertaling
23	20:33 Operator: maar het is belangrijk om daarover te blijven praten, onder andere **met je psycholoog**	mais il est important de continuer à en parler, y compris avec votre psychologue **Letterlijke vertaling** maar het is belangrijk om hierover te blijven praten, met inbegrip van uw psycholoog
	Chatregel	**Machinevertaling**
24	20:33 Meisje: normaal zou ik ook niet naar hier, maar in de plaats naar haar gaan, maar ik zie haar pas over een maand	normalement, je n'irais pas ici non plus, mais plutôt chez elle, mais je ne la verrai pas avant un mois. **Letterlijke vertaling** normaal zou ik hier niet meer naartoe gaan maar eerder naar haar, maar ik zal haar pas over een maand zien.
	Chatregel	**Machinevertaling**
25	20:34 Operator: **die kan kaderen waar het trauma vandaan komt en hoe je er dagdagelijks mee om kan gaan.** mensen die geen trauma hebben meegemaakt, kunnen zich soms moeilijker blijven inleven in hoe iemand zich jaren nadien nog voelt	**[die kan kaderen waar het trauma vandaan komt en hoe je er dagdagelijks mee om kan gaan].** Les personnes qui n'ont pas vécu de traumatisme ont parfois du mal à continuer à ressentir de l'empathie pour les sentiments d'une personne des années plus tard. **Letterlijke vertaling** Mensen die geen trauma hebben meegemaakt vinden het som moeilijk om empathie te blijven voelen voor iemands gevoelens jaren later.

CLB6-R23-25, Frans, impactniveau major

- Ook **losse woorden in een chatregel leiden tot problematische vertalingen**. Wanneer de hulpverlener de cliënt een compliment wilde geven voor een bepaalde actie en alleen het woord "Top!" intypte in de chat, werd dit in alle talen door DeepL vertaald in de betekenis van "bergtop" in plaats van met een equivalente term met de juiste betekenis (bijv. 'goed gedaan!').
- Wanneer **er weinig interpunctie** (komma's, punten etc.) wordt gebruikt, kan dit in de MT-output leiden tot **fouten in de zinsstructuur en grammaticale problemen** (zo kan bijvoorbeeld het onderwerp bij een werkwoord veranderen in de MT).
- Veel **spelfouten** worden automatisch gecorrigeerd in de MT, maar er zijn uitzonderingen. In sommige gevallen wordt het onjuist gespelde woord **letterlijk overgenomen in de MT**, wat **verwarring** op kan leveren wat betreft de betekenis (bijv. "schoom" i.p.v. "school", dat dan letterlijk wordt overgenomen in de vreemde taal).

3.3 Gebruikerservaringen

- Medewerkers die deelnamen aan de livegesprekken hebben die overwegend als **positief en vlot** ervaren, waarbij zij aangaven dat dit mede zo was dankzij de welwillende houding van de cliënten.
- Ondanks dat ze de brontaal niet kenden, hadden ze de indruk dat de vertalingen **redelijk accuraat** waren en **dat de cliënten hen goed begrepen**.
- Eén medewerker liet de logs nalezen door een collega met kennis van Turks, die wees op het **formele taalgebruik**. Deze instelling kan echter aangepast worden in DeepL (formeel/informeel taalgebruik).
- Wel werden **verschillende vertaalfouten** opgemerkt, vooral in het Arabisch, waarin DeepL gebruikt soms "ongepaste termen" suggereerde.
- Medewerkers wijzen er dan ook op de vertaaltechnologie **minder geschikt is voor emotioneel beladen gesprekken** en beter voor eerder 'lichtere' gesprekken ingezet zou moeten worden.
- Meerdere medewerkers gaven aan dat de **huidige methode van kopiëren en plakken via DeepL omslachtig is en vertragend** werkt. Het constant switchen tussen vensters (chatvenster en DeepL) werpt een technische drempel op en verstoort het gesprek, waardoor medewerkers het moeilijk vinden om inhoudelijk bij te blijven. Ze vermelden dan ook dat het belangrijk is om op elkaar te wachten en niet door elkaar heen te chatten.
- Alle medewerkers die deelnamen aan de gesprekken zouden **bereid zouden zijn vertaaltechnlogie te gebruiken** in hun chathulpverlening indien **directe vertaling** van de chats (via API) mogelijk zou zijn.
- Het gebruik van vertaaltechnologie wordt door iedereen als een **meerwaarde in hun communicatie** met anderstalige cliënten beschouwd vanwege de toegankelijkheid en snelheid (de inzet van een vertaaltool als **laagdrempelig alternatief** in situaties waarin het lastig is om een tolk te regelen), **de taalnabijheid voor de cliënt** (cliënten waardeerden het om in hun eigen taal te kunnen spreken) en de **menselijke connectie om via emoji's en informele toon** het gevoel van nabijheid te kunnen vergroten.
- Een probleem dat soms werd vermeld bij de livegesprekken was het **regelmatig wegvallen van de internetverbinding**. Daardoor moesten soms andere communicatiemiddelen ingezet worden zoals Whatsapp.

4. Strategieën op microniveau

Op basis van de onderzoeksresultaten kunnen strategieën worden geformuleerd die medewerkers kunnen toepassen om MT op de minst risicovolle manier te kunnen gebruiken. Hiervoor kunnen de voorbeelden die zijn voorgesteld in dit rapport als uitgangspunt dienen. Zo kunnen medewerkers erop letten dat ze hun boodschap op een heldere, niet-dubbelzinnige manier formuleren, de juiste interpunctie (komma's en punten) gebruiken en dialectwoorden zoveel mogelijk vermijden. De chattekst moet dus geoptimaliseerd worden, om het risico op een onjuiste vertaling te beperken. De analyses van de livegesprekken laten zien dat in het geval van een onduidelijkheid, al dan niet veroorzaakt door de MT, de hulpverlener een bijkomende vraag kan stellen om zeker te zijn dat de boodschap correct is overgekomen. Aangezien van de hulpvragers slechts in beperkte mate verwacht kan worden dat zij hun chatteksten optimaliseren, moeten hulpverleners er extra op attent zijn dat de MT hiervan onjuistheden kan bevatten. Maar ook hulpvragers zouden (beperkte) instructies moeten krijgen bij het gebruik van MT. Voor specifieke, hands-on oplossingen, zie paragraaf 6 (Webinar en leermiddelen). Een andere strategie gericht op risicomanagement bestaat uit het opstellen en integreren van glossaria met vaste vertalingen (of eenvoudigweg het handhaven van) veelvoorkomende afkortingen, acroniemen en gespecialiseerde terminologie in de vertaalsoftware.

5. Conclusies

5.1 Algemene bevindingen

- Bij **anonieme chatgesprekken** bestaat een risico op miscommunicatie door een gebrek aan context en informatie over de cliënt, terwijl in de **blended-chatgesprekken** is het risico van miscommunicatie kleiner is, omdat hier met gekende cliënten wordt gewerkt (bij wie de hulpverleners ook in een bepaalde mate inzicht hebben in hun problematiek) en eventuele misverstanden die het gevolg zijn van een onjuiste vertaling wellicht makkelijker worden gesignaleerd dan bij anonieme chathulpverlening.
- Medewerkers geven daarnaast zelf aan dat **zij MT geschikt achten voor gesprekken die eerder van "lichtere" inhoudelijk of praktisch karakter zijn** en **minder voor emotioneel beladen** gesprekken, niet alleen vanwege de **kwaliteit van de vertaling** maar ook vanwege de **vertraging** in de uitwisselingen.
- Het valt op dat in de livegesprekken, **hulpverleners sneller reageren** op opvallend 'vreemde' vertalingen, terwijl het voorkwam in de gesprekken dat cliënten niet reageerden op vertalingen die voor veel taalgebruikers als overduidelijk onjuist zouden worden gepercipieerd. Daarom is het zeer belangrijk dat hulpverleners extra alert zijn op eventuele misverstanden. Ook blijft het de vraag in hoeverre deelnemers zich bewust (kunnen) zijn van bepaalde vertaalfouten. Wel moet rekening gehouden worden met het feit **dat verschillende deelnemende cliënten het Nederlands deels machtig zijn** en de beide talen zagen staan in de chat (de brontekst en de MT), wat hun begrip van de tekst wellicht vergemakkelijkte, waardoor zij geen reactie vertoonden.
- Het gebruik van MT brengt dus een **nog grotere verantwoordelijkheid voor de hulpverleners** met zich mee dan eentalige communicatie: zij moeten extra aandacht hebben voor hun eigen taalgebruik en het communicatieproces nog **scherper monitoren** dan zij in eentalige communicatie al doen.
- Het belangrijkste doel op dit moment bij het gebruik van MT is **bewustwording** bij de medewerkers **van mogelijke risico's** bij het inzetten van digitale vertaalmiddelen het anticiperen

op en het voorkomen van miscommunicatie door middel van een duidelijk geformuleerde en heldere communicatiestijl.

- Afgezien van wat de medewerkers zelf kunnen doen, moeten **instellingen** zorgen voor **beveiligde accounts** wanneer zij een vorm van MT of AI gebruiken en gebruikersovereenkomsten met cliënten om hun dataveiligheid te garanderen.
- Ideaal gezien zouden in die betaalde accounts ook **glossaria** worden opgenomen voor gangbare en belangrijke termen en afkortingen binnen de gebruikte jargons, zodat deze op consistente wijze worden vertaald.
- **Integratie van MT** in de **chatsoftwaretool via API** zou een **gebruikersvriendelijkere**, **efficiëntere** en **snellere communicatie** opleveren, en wanneer daarbij ook AI wordt toegepast, zou het risico of fouten die worden veroorzaakt door **een gebrek aan context** kunnen **verminderen**. Verder **empirisch onderzoek** op basis van authentieke data van partners uit de praktijk is nodig op dit gebied.
- Instellingen moeten daarnaast **duidelijke richtlijnen ontwikkelen** voor een verantwoord gebruik van digitale middelen in de communicatie met anderstalige hulpvragers.
- Op een hoger beleidsniveau (Vlaamse regering) is er dringend **wetgeving** nodig op het gebied van **digitale communicatiemiddelen en vertaaltechnologie**.

5.2 Bredere tendensen op het gebied van de inzet van digitale middelen in hulpverlening/openbare diensten

- Er is een grote **behoefte aan informatievoorziening over het gebruik van MT/AI** bij medewerkers van openbare dienstverleningsinstellingen, die in hun dagelijkse communicatie met hulpvragers frequent te maken krijgen met anderstaligen. Medewerkers willen weten of **zij mogen communiceren** door middel van MT/AI en **hoe** ze dat het best kunnen doen.
- Belangrijk is ook dat er **nauwelijks alternatieven zijn voor MT** in de chathulpverlening (behalve het gebruik van een lingua franca zoals het Engels) wanneer hulpverleners cliënten willen helpen in een andere taal dan het Nederlands, omdat er **geen beroep** kan worden gedaan op **een professionele vertaler of tolk.** Dat wil zeggen dat cliënten die geen Nederlands spreken in principe **alleen toegang** tot deze dienst kunnen krijgen **door middel van vertaaltechnologie**. Om **sociale gelijkheid te bevorderen** is de inzet van vertaaltechnologie dus belangrijk.
- Op dit moment zijn er **geen algemene richtlijnen** vanuit die instellingen of vanuit de overheid met betrekking tot de inzet van machinevertaling en AI.
- Medewerkers proberen om te kunnen communiceren met anderstaligen **toch bepaalde middelen in te zetten.**
- Die middelen zijn vaak **gratis**, waardoor de **dataveiligheid** van cliënten (en medewerkers) ernstig in gevaar komt.
- Dit is in strijd met het **dataveiligheidsbeleid** van openbare instellingen in Vlaanderen, terwijl de chatsoftware (SIT-tool) zelf wel aan zeer strenge dataveiligheidseisen moet voldoen.
- Een **veilige, beschermde digitale omgeving** waarin **MT/AI lokaal wordt ingezet** zou een oplossing voor dit probleem kunnen bieden.

5.3 Beperkingen van het onderzoek

- De resultaten van dit onderzoek kunnen enigszins **vertekend** zijn omdat een groot deel van de analyses is uitgevoerd op **bestaande, statische transcripten** van anonieme chatgesprekken. Aangezien deze data statisch is, kan niet worden nagegaan of misverstanden die zijn ontstaan

naar aanleiding van onjuiste vertalingen (deels) zouden zijn opgelost. Dit in tegenstelling tot de livegesprekken, waarin onduidelijkheden of misverstanden tijdens het gesprek in principe wel konden worden opgehelderd.

- De transcripten van de anonieme (statische) chatgesprekken waren volledig in het Nederlands en werden vertaald via DeepL naar het Arabisch, Turks en Frans, wat een enigszins **kunstmatige methode is**, aangezien de hulpvragers in een livegesprek hun deel van de chat in de eigen taal zouden intypen en niet in het Nederlands. Het geeft echter wel een beeld van de vertaalproblemen die in de onderzochte talencombinaties kunnen optreden. Ook hebben we bij het kiezen van voorbeelden ter illustratie van fouten vooral chatfragmenten geselecteerd waarin de tekst van de hulpverlener werd vertaald naar de vreemde taal, omdat deze de realiteit van de chathulpverlening het beste weerspiegelen.
- **De resultaten van het onderzoek zijn gebaseerd op de machinevertalingen van één specifieke vertaalsoftware**, namelijk DeepL, die hoog staat aangeschreven in vertaalopleidingen en bij professionele vertalers. Desondanks moet steeds in beschouwing worden genomen dat het gebruik van andere vertaalsoftware (bijvoorbeeld Google Translate) tot nog andere resultaten zou kunnen leiden. De vergelijking van verschillende softwareprogramma's maakte geen deel uit van dit onderzoek.
- Door **vertraging in het aanleveren van de data en in de revisie** door de vertalers was de dynamische dataset kleiner dan voorzien (8 bruikbare gesprekken in plaats van 20) en kon **niet de volledige dataset** worden geanalyseerd.
- Gezien deze beperkingen kunnen de resultaten van de analyse **niet gegeneraliseerd** worden. Het onderzoek dient daarom ook als **exploratief** te worden beschouwd en vormt hopelijk de aanzet tot een grootschaliger en breder onderzoek naar de inzet van digitale vertaalhulpmiddelen in de online en blended hulpverlening in Vlaanderen.
- In lijn met de doelstelling van het huidige onderzoek, kan toekomstig onderzoek zich richten op het **analyseren van vergelijkbare, grotere datasets** waarbij een **breder scala aan talen** wordt onderzocht. Met het oog op de snelle ontwikkelingen op het gebied van **generatieve artificiële intelligentie (GenAI)** is het bovendien zeer nuttig om te onderzoeken in hoeverre het gebruik van "large language models", die veel breder gaan dan de neurale netwerken van vertaaltechnologie zoals DeepL en daardoor ook meer context voor tekstinterpretatie kunnen integreren, kunnen leiden tot een **verbetering van de kwaliteit** van automatische vertalingen.

6. Webinar en leermiddelen

- **Opzet en inhoud webinar**
    - Hulpverleners bevragen over hun ervaringen met chat met anderstaligen en het gebruik van vertaaltechnologie in chats met anderstaligen
    - Delen van de onderzoeksresultaten
    - Overzicht geven van de meerwaarde en valkuilen van het gebruik van vertaaltechnologie in chatgesprekken
    - Aanbevelingen geven aan organisaties die vertaaltechnologie (willen) gebruiken in chatgesprekken. Dit op vlak van veiligheid van chat- en vertaalsoftware.
    - Richtlijnen geven aan hulpverleners die vertaaltechnologie (willen) gebruiken. Dit op inhoudelijk en technisch vlak.
    - Uitwisseling tussen de deelnemers over hun ervaringen met vertaaltechnologie in chatgesprekken.

    - Hulpverleners een AI-tool aanbieden waarmee ze kunnen oefenen met het gebruik van vertaaltechnologie in chatgesprekken.
- **Leer- en communicatiemiddelen**
    - Handleiding voor het gebruik van vertaalsoftware in chatgesprekken
    - Oefenen met vertaalsoftware in chatgesprekken
    - ChatGPT - Meertalige Chat Trainer SAM (te gebruiken met de “Handleiding” en het document “Oefenen met vertaalsoftware in chatgesprekken”)

Plan Overview

A Data Management Plan created using DMPonline.be

Title: NEW_"CHAT SUPPORT USING TRANSLATION TECHNOLOGY" (CHATT) – VERTAALTECHNOLOGIE VOOR CHATHULPVERLENING

Creators: Esther de Boe, n.n. n.n., Catherine Meulemans

Affiliation: University of Antwerp

Template: BOF Template (Flemish Standard DMP)

Principal Investigator: Esther de Boe

Data Manager: n.n. n.n.

Project abstract:

Heel wat organisaties in Vlaanderen, waaronder CAW en CLB, bieden actief chathulp aan, zowel met gekende cliënten op afspraak als via een anonieme chattool. De organisaties worden steeds vaker geconfronteerd met cliënten, die niet (voldoende) kunnen geholpen worden in het Nederlands. Omdat in chathulpverlening geen beroep kan gedaan worden op sociaal tolken, zoeken CAW en CLB naar mogelijkheden om vertaaltechnologie in te schakelen in de chat om zo dit kanaal ook toegankelijker te maken voor deze cliënten. Met de eerste fase van het project voeren we een wetenschappelijk onderzoek naar de kwaliteit van machinevertalingen van hulpverleningsgesprekken. De onderzoeksvraag die we daarbij willen hanteren is de volgende: In welke mate kan machinevertaling via DeepL emotioneel geladen gesprekken in realtime chathulpverlening ondersteunen? In de tweede fase zullen CAW en CLB inzetten op de ontwikkeling van een professioneel handelingskader en op de ontwikkeling van een gepast trainingsaanbod voor het gebruik van vertaaltechnologie in chathulpverlening. Met het project willen we een eerste stap zetten in de richting van het integreren van vertaaltechnologie in chatsoftware van CAW en CLB. De onderzoeksresultaten en het ontwikkelde handelingskader zullen daar in belangrijke mate aan bijdragen.

ID: 210754

Start date: 20-11-2023

End date: 09-12-2025

Last modified: 05-12-2024

NEW_“CHAT SUPPORT USING TRANSLATION TECHNOLOGY” (CHATT) – VERTAALTECHNOLOGIE VOOR CHATHULPVERLENING

Project Information

Antigoon project ID

53452

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
1. Bestaande transcripten	20 geanonimiseerde fragmenten van transcripten van bestaande chatgesprekken, in de originele Nederlandse versie én een revisie van deze transcripten in het Arabisch, Turks en Frans door professionele vertalers.	Reuse existing data	Digital	Observational	txt	<1GB	
2. Realtime transcripten	20 geanonimiseerde fragmenten van transcripten van realtime chatgesprekken (gemiddelde duur 15 minuten) van CLB-/CAW-medewerkers met anderstalige cliënten+ revisie van deze transcripten.	Generate new data		Observational	txt	<1GB	
3. Enquête	De geanonimiseerde data van de beknopte tevredenheidsenquêtes die de medewerkers na afloop van de gesprekken invullen. Deze worden opgesteld door de UAntwerpen via Qualtrics en digitaal uitgevoerd door de medewerkers van CAW/CLB.	generate new data		Observational	txt	<1GB	
4. Informed consent + infoblad	De medewerkers en cliënten geven hun toestemming door middel van het aankruisen van een hokje voor akkoord, cliënten krijgen een versie vertaald in de eigen taal. Papieren versies worden ingescand en bewaard door CLB/CAW. UAntwerpen heeft geen toegang tot deze data, ontvangt alleen de bevestiging van de partners dat er i.c. is verkregen.	new data		Observational	txt	< 1GB	
5. Rapportage van de resultaten	Overzicht van de resultaten	new data		Observational	txt	<1GB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

De geanonimiseerde fragmenten van transcripten van de bestaande data (tekstbestanden) en de revisies hiervan in het Turks en het Arabisch worden aangeleverd door de partners CLB/CAW. Voor de analyses van de transcripten van de bestaande gesprekken kan achteraf geen toestemming worden gevraagd omdat een groot deel hiervan anoniem is gevoerd, dus met cliënten die niet bekend zijn bij de instellingen. Dit in tegenstelling tot de realtimegesprekken, die uitsluitend worden gehouden met gekende cliënten in het kader van de blended hulpverlening. Het gebruik van de transcripten van de bestaande gesprekken is te rechtvaardigen omdat deze opnieuw in geen enkel opzicht te herleiden zullen zijn tot de betrokken personen. Bovendien gaat het om fragmenten uit gesprekken, wat ook bijdraagt aan niet-identificatie.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, human subject data

Het betreft maximaal 36 deelnemers: 6 medewerkers CAW + 12 cliënten/6 medewerkers CLB + 12 cliënten. Deelnemers zijn tussen de 13 en 65 jaar. De 12 medewerkers van CAW/CLB zijn Nederlandstalig en betrokken bij blended chathulpverlening, waarbij een deel van de zorgverlening plaatsvindt via contact ter plaatse en een deel via realtime chat; de 12 cliënten zijn anderstalig (Frans/Arabisch of Turks) en zijn cliënten van de genoemde instellingen die gebruik maken van blended hulpverlening, waarbij een deel van de zorgverlening plaatsvindt via contact ter plaatse en een deel via realtime chat. Alle deelnemers ontvangen een informatieblad en geven hun akkoord in een informed-consentformulier (de anderstaligen in hun eigen taal). Alle deelnemers worden gerekruteerd via de instellingen CAW en CLB.

Aangezien de cliënten over het algemeen een hulpvraag hebben als zij bij het CAW/CLB aankloppen, is de waarschijnlijkheid groot dat zij als kwetsbaar kunnen worden beschouwd. De data die uit de gesprekken wordt verzameld blijft echter volledig anoniem. Daarnaast betreft het communicatie met medewerkers die gespecialiseerd zijn in het bieden van hulp aan kwetsbare personen, die er zeker oog voor zullen hebben als er zich een risico voor zou doen, aangezien dat deel uitmaakt van hun dagelijkse werk.

De thema's die in de gesprekken aan bod komen kunnen zeker gevoelig zijn, aangezien cliënten van het CAW en CLB daar komen met een vraag om hulp, in welke vorm dan ook. De vertaling van deze gevoelige onderwerpen en evt. emoties die deze oproepen bij de cliënten zijn een belangrijk onderdeel van het onderzoek. Zie verder hierboven: de cliënten zijn bij de medewerkers in goede handen juist met betrekking tot deze gevoelige onderwerpen.

Eventuele ongemakken voor de deelnemers zouden kunnen veroorzaakt worden wanneer de machinevertaling niet juist zou zijn en tot onbegrip tussen de 2 partijen zou leiden. Ook hier kan als oplossing aangehaald worden dat de medewerkers ruime training hebben gehad in het oplossen van psychosociale problemen en veel ervaring hebben in communicatie met deze groepen. Ze kunnen evt. doorverwezen worden door de hulpverleners naar gespecialiseerde instellingen. Momenteel is er geen hulp mogelijk in hun moedertaal, waardoor er in de huidige situatie geen optimale communicatie is. Daarnaast is het zo dat tijdens het gesprek niet enkel in de vertaalde taal wordt gecommuniceerd, maar ook in het Nederlands, aangezien de medewerkers steeds de tekst in het Nederlands én de vertaalde tekst in de chat ingeven. De vertaling is dus een toegevoegde waarde op de bestaande communicatie.

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes

Alle persoonlijke elementen worden door CLB/CAW verwijderd alvorens te worden gedeeld met de UAntwerpen.

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Dit project omvat een verkennend onderzoek waarbij we de kwaliteit van bestaande machinevertaling voor chathulpverlening nagaan, maar zelf niets nieuws ontwikkelen. Valorisatie van de onderzoeksresultaten van het huidige project bestaat vooral uit het formuleren van adviezen naar de partners CAW/CLB toe over het gebruik van machinevertaling voor chathulp.

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- Yes

Alle aangeleverde transcripten en revisies vallen onder de geheimhoudingsplicht.

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

Er wordt een map gecreëerd in SharePoint met als naam: "CHATT". Die map wordt onderverdeeld in verschillende mappen (Map 1 -Dataset 1 = Transcripten; Map 2 - Dataset 2 = Informatieblad/Lege Informed consentformulieren; Map 3 - Dataset 3 = Survey; Map 4 - Dataset 4 = Output); Map 5 Ethical clearance, waarin de desbetreffende data worden ondergebracht.
Map 1 (Submappen)

- Bestaande chat (Brontekst; Professionele revisie; Machinevertaling; Analyses)
- Realtime chat (Brontekst; Professionele revsie; Machinevertaling; Analyses)

Map 3:
Survey

- Antwoorden
- Analyse

Map 4:

- Academisch (wetenschappelijke papers)
- Niet-academisch (leermiddelen CLB/CLW

Map 5:

- Ethische commissie
- DMP
- Overeenkomsten
- Vertouwelijkheidsverklaring
- Geheimhoudingsverklaring studenten (deelnemende MA-studenten die de data gebruiken voor hun scriptie)

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

3. Data storage & back-up during the research project

Where will the data be stored?

During the course of the project, the PI will be responsible for the preservation of the data in an individual folder on the Microsoft SharePoint Sites of the Department Applied Linguistics/Translators and Interpreters. The folder will be accessible only to the PI and the research manager of the Faculty of Arts.

How will the data be backed up?

The Microsoft SharePoint Site automatically backs up content every 12 hours and a retention policy protects all data in the SharePoint Site of the Departments for 5 years. This policy allows the website administrators, Géraldine De Visscher, to manually retain items for 5 years based on when they were last modified. In addition, Microsoft guaranties that all data will be mirrored in

multiple data centres and completely versioned.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

SharePoint Online has a storage capacity of 25 TB per SPO site.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The individual folders on the Departmental SharePoint Site will be accessible only to the PI
and by the research manager of the Faculty of Arts.
Write access to these SharePoint folders is based on authentication via a personal username and password. For members of the Faculty of Arts of the University of Antwerp, these usernames and passwords are managed by the UA ICT Department (Active Directory). For co-supervisors linked to another university or research institution, the research manager will send a link to the folder via email, and as such give them write access as well. The link will be usable only by the owner of the email account.
All data changes are traceable and all files are versioned. Changed files can be reversed to a previous version and accidentally deleted files can be recovered from the SharePoint trash bin.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

All eventual (extra) costs will be covered by the University of Antwerp, and are not included in the project budget.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All data.

Where will these data be archived (stored and curated for the long-term)?

All data will be archived for the required retention period of 10 years in the mentioned folders, managed by the research manager of the Faculty or Arts.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

All eventual (extra) costs will be covered by the University of Antwerp, and are not included in the project budget.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- No (closed access)

If access is restricted, please specify who will be able to access the data and under what conditions.

Esther de Boe heeft toegang tot alle data om deze te gebruiken voor onderzoeksdoeleinden. Daarnaast hebben 2 masterstudenten toegang tot de transcripten voor de duur van hun onderzoek, waarvoor zij een geheimhoudingsplicht en afstand van rechten hebben ondertekend. Dit geldt ook voor de onderzoeksmedewerker die gedurende 2 maanden de resultaten gaat analyseren. De resultaten van de analyses worden gedeeld met de partners in de vorm van een rapportage van de eindresultaten. Deze kunnen de partners gebruiken voor de ontwikkeling van leermiddelen voor hun medewerkers. De ontwikkeling hiervan staat los van de de activiteiten van de UAntwerpen en wordt volledig uitgevoerd door de partners.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

-

When will the data be made available?

Bij publicatie van de resultaten.

Which data usage licenses are you going to provide? If none, please explain why.

N.v.t.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

What are the expected costs for data sharing? How will these costs be covered?

N.v.t.

6. Roles & Responsibilities

Who will manage data documentation and metadata during the research project?

Esther de Boe

Who will manage data storage and backup during the research project?

Esther de Boe (p.i.)

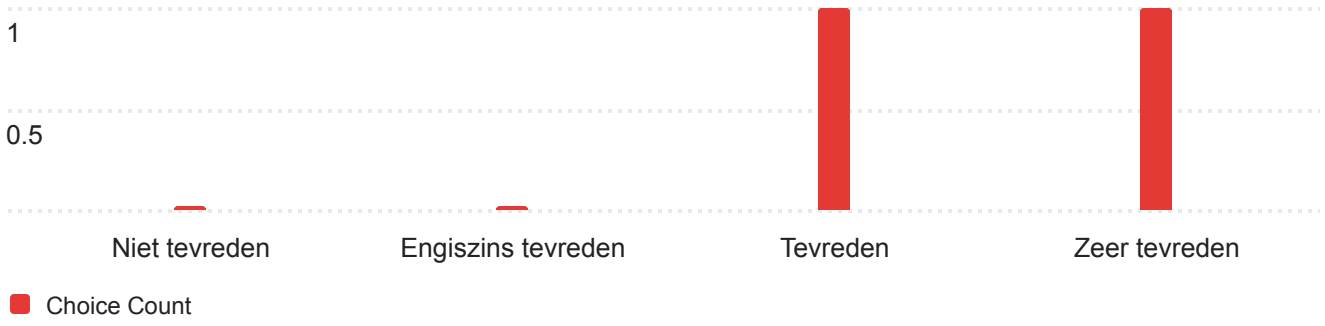
Who will manage data preservation and sharing?

Esther de Boe

Who will update and implement this DMP?

Esther de Boe

Q1 - 1. Hoe tevreden ben je over de vertaling?



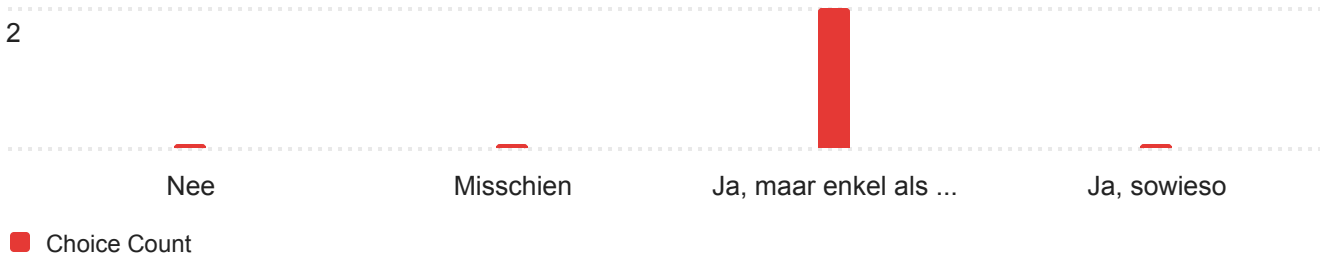
Q2 - 2. Heb je bijkomende vragen moeten stellen of verduidelijking moeten vragen door problemen met de vertaling?

Q3_1 - Helemaal niet

Field	Min	Max	Mean	Standard Deviation	Variance	Responses	Sum
Helemaal niet	65.00	95.00	80.00	15.00	225.00	2	160.00

Q4 - 4. Hoe gemakkelijk of moeilijk is het kopiëren en plakken van de vertaling?

Q5 - 5. Zou je deze vertaaltechnologie in de toekomst vaker willen gebruiken?



Q6 - 6. Heb je opmerkingen of suggesties? Of heb je bijvoorbeeld feedback ontvangen van de cliënt met betrekking tot het verloop van het gesprek en over het gebruik van de vertaaltechnologie? Dan mag je die hier intypen. Alvast bedankt voor je medewerking!

6. Heb je opmerkingen of suggesties? Of heb je bijvoorbeeld feedback ontvangen van de cliënt met betrekking tot het verloop van het gesprek en over het gebruik van de vertaaltechnologie? Dan mag je die hier intypen. Alvast bedankt voor je medewerking!

Het vertalen, kopiëren, plakken, antwoorden verliep traag. Dit viel ook op voor de cliënt. Doordat ik niet in het gesprek kon gaan terugkijken naar vorige berichten (omdat ik het niet versta), maar daarvoor op Deepl moest berusten, is het ook onoverzichtelijk welke berichten door wie zijn geschreven. Dit neemt ook weer tijd in beslag. Je moet er ook aan denken om de vorige tekst te behouden in het Deepl venster.

De taal was Arabisch, dus het selecteren om te copy pasten ging van rechts naar links. Dit was verwarrend en leek soms ook niet te lukken en dat is wat frustrerend. Vooral omdat je cliënt een antwoord verwacht, en dit was een emotioneel geladen gesprek.

De inhoud van de vertaling was wel helemaal ok en de cliënt heeft meer kunnen zeggen omdat het in haar eigen taal was.

Een suggestie is om de vertaalsoftware te integreren bv door de NL vertaling vlak langs de Arabisch weer te geven. En in je eigen vak iets te typen wat dan automatisch vertaald wordt in het chatvenster

Indien geïntegreerd zou dit enorme meerwaarde zijn voor vlotte gesprekken online. Client leek ook hele tevreden over Turkse vertaling.

Oproep medewerking aan het project 'Inzetten van vertaaltechnologie in chathulpverlening'.

In juni 2024 dienden CAW en CLB tezamen met de Universiteit Antwerpen een projectaanvraag met als titel 'Inzetten van vertaaltechnologie in chathulpverlening' in bij het Departement Zorg. Een subsidie werd goedgekeurd, met het oog op het uitvoeren van een onderzoeksproject in de loop van het schooljaar 2024-2025.

Om het project te kunnen realiseren hebben we jou nodig. Door mee te werken draag je actief bij aan de doelstelling van het project: Het onderzoeken van de mogelijkheden om vertaaltechnologie in te schakelen in de chathulpverlening van CAW en CLB om dit kanaal toegankelijker te maken voor cliënten die het Nederlands niet of minder machtig zijn.

Hieronder geven we een overzicht van de verwachtingen en werkwijze.

VRAAG	ANTWOORD
Wie zoeken we?	- 6 CAW-medewerkers en 6 CLB-medewerkers die bereid zijn om mee te werken aan het onderzoeksproject.
Wat wordt er verwacht van jou als hulpverlener?	- Tussen november 2024 en uiterlijk 31/1/2025 (minstens) **2 chatgesprekken voeren** met cliënten (van 13 jaar of ouder) die het Nederlands niet of onvoldoende machtig zijn en die als moedertaal het Standaard Arabisch, Turks of Frans hebben. - Je kruist hiervoor (eenmalig) een 'informed consent' formulier aan voor akkoord, waarmee je aangeeft dat je bereid bent om aan het onderzoek mee te werken. - De chatgesprekken worden gevoerd via de SIT-tool (CAW-medewerkers en CLBch@t medewerkers die toegang tot de SIT-tool hebben) of via de PHP-live software voor blended chatten (andere CLB-medewerkers). - Met het oog op de veiligheid van de data, gebruik je een betalende account van DeepL op jouw naam. De accountgegevens worden je bezorgd van zodra je toegezegd hebt om deel te nemen aan het onderzoek. - De chatgesprekken bezorg je ons geanonimiseerd (zie onder).
Hoe verloopt het chatgesprek?	- Als hulpverlener begin je met het typen van een zin in het Nederlands, rechtstreeks in DeepL (vertaalsoftware). De NL zin + vertaling kopieer je, plak je in het chatvenster van de SIT-tool/PHP Live en verstuur je in 1 geheel naar de cliënt. - De cliënt typt zijn chatbericht in de eigen moedertaal. Je kopieert dat chatbericht en plakt dat in DeepL. Op basis van die vertaling reageer je weer in het Nederlands én in de moedertaal van de cliënt. - Dit werkt het makkelijkst met een de splitscreen functie (bv links de chatsoftware en rechts DeepL), of door met 2 schermen te werken. - Dat ziet er dan bijvoorbeeld als volgt uit (zie onderstaande printscreen uit de SIT-tool van een fictief gesprek met een Arabisch sprekende cliënt).

VRAAG	ANTWOORD
	 - Na afloop van het gesprek download je het chattranscript en anonimiseer je het (zie instructies over het anonimiseren van een chattranscript en zie demovideo). - Na het voeren van je 2 gesprekken vul je als hulpverlener nog een kort evaluatieformulier in.

Hoe informeer ik de cliënt over medewerking aan het onderzoek?	- We bezorgen jou 2 documenten die zowel in het Nederlands als in de taal van de cliënt opgesteld zijn.
	  ○ Een informatieblad, waarin de bedoeling van het onderzoek is omschreven: Wat er verwacht wordt en waarom we dit organiseren.
	  ○ Een 'informed consent'- formulier, dat (digitaal of op papier) door de cliënt voor akkoord moet aangekruist worden worden.

Waaraan moet een gesprek voldoen?	- Het chatgesprek moet minstens 15 minuten duren. - Het moet een chatgesprek zijn waarin inhoudelijke (mogelijk emotioneel geladen) thema's besproken worden. Het mag dus niet enkel gaan over bijvoorbeeld het maken van een afspraak of het geven van informatie.

Praktisch:

- Ben je CAW-medewerker en wens je deel te nemen aan het onderzoek? Neem dan contact op met berbe.luyckx@cawgroep.be.
- Ben je CLB-medewerker en wens je deel te nemen aan het onderzoek? Neem dan contact op met tom.billiet@clbonline.be.

Twijfel je nog, dan kan je uiteraard via bovenstaande mailadressen ook terecht voor meer info.

Richtlijnen revisie fragmenten chatgesprekken CLB/CAW

Beste vertaler,

In de linkerkolom zie je de originele tekst en in de rechterkolom de machinevertaling door DeepL. Wanneer je een probleem tegenkomt in de machinevertaling (zie de indicaties hieronder), volg dan deze stappen:

- Geef in de linker én rechterkolom aan waar de fout zit door middel van onderlijnen of **markeren in vet**.
- Laat in de linker kolom in de regel direct onder de brontekst zien wat de verbetering is (de **juiste** vertaling van het foute element).
- Laat in de rechterkolom direct onder de machinevertaling zien wat **fout** is door middel van een letterlijke vertaling van de fout.
- Indien nodig kun je toelichting geven in de vorm van een opmerking/ballonnetje (lint: controleren → nieuwe opmerking). Als de fout duidelijk blijkt uit de letterlijke vertaling, is dit niet nodig

Mogelijke fouten:

- Weglating van een woord of zinsdeel
- Onjuiste toevoeging
- Onjuiste vertaling
- Onjuist taalgebruik
- Onjuiste spelling
- Onjuiste grammatica
- Onjuiste terminologie
- Stijlfout (bijv. “u” in plaats van “jij”)
- Woord/tekst onvertaald overgenomen

Voorbeelden revisie

Kun je daar iets meer over vertellen?	¿Puede contar algo más sobre eso?
¿Puedes	*Kunt u*
Ik praat niet tegen mijn stief papa, hij ook niet tegen mij, mijn zus is vaak op kot.	No hablo con mi padrastro, él tampoco habla conmigo, mi hermana suele estar en casa
en su cuarto de estudiante	
Je woont dus in een nieuw samengesteld gezin	Así que estás viviendo en una **nueva** familia ensamblada
una familia ensamblada	
Met mijn stiefpapa praat ik niet.	con mi padrastro no hablo con nadie.
	met mijn stiefpapa praat ik niet met niemand.
Of is dit de eerste keer dat je met iemand praat over je thuissituatie?	¿O es la primera vez que hablas con alguien sobre la situación **de tu** hogar?
tu situación de hogar?	*de situatie van jouw huis?*

NB: De kleuren hebben geen betekenis, deze zijn ter verduidelijking van de eenheden vertaling + revisie.

Hartelijk dank voor je medewerking! Heb je vragen? Mail: esther.deboe@uantwerpen.be