



ATA GLD Members in
Europe
9th Annual Workshop
February 2-4, 2024

The New Kid on the Block: First Experiences with GenAI in Translation

Dipl.-Dolm. Johanna Klemm, JK Language Services Europe LTD,

ATA-certified EN-DE, Kenilworth, UK

About me

- Dipl.-Dolm. Johanna Klemm, JK Language Services Europe LTD
- ATA-certified EN-DE, Kenilworth, UK
- Freelance Translator and Conference Interpreter for over 20 years
- German (native), English, French, Portuguese
- Medical, marketing, business
- Several major technology shifts during my professional life – Internet, TM, dictation software, (N)MT and now AI?

Goals

- Basic concepts and questions
- Forms of AI use for translation
- Concrete examples for AI in a CAT tool
- Value of the human translator
- Discussion

To use AI or not to use AI?



By ESA/Hubble, CC BY 4.0,
<https://commons.wikimedia.org/w/index.php?curid=8788068>

(N)MT vs. GenAI

(N)MT

- Mature technology
- Purpose:
 - Transforms text between languages
 - Specializes in accurate language conversion
- Faster, uses fewer resources
- Can translate long documents
- Is static
- Is part of an environment designed for the translation workflow
- Generates professional deliverables (xliff files)

GenAI

- New technology
- Purpose:
 - For conversational interaction, all-purpose tool
 - Designed for versatile text generation (context-oriented)
- Slower, uses more processing power
- Has token limitations
- Can generate different translations/fine-tune text, is trainable
- Has a general-purpose environment/chat interface
- Generates only „simple“ text or tables

Developed in conversation with the organizers and participants of the techforwords AI Bootcamp January

Uses of GenAI in translation

- Support tasks
- AI for translation in novel environments
- AI embedded in a CAT tool

Support tasks (not in a CAT tool)

Examples:

- Prepare source text
- Research – definitions, explanations
- Identifying terminology, glossaries
- Alternatives, synonyms, metaphors
- Fine-tuning translation

To do:

- Explore different options, ChatGPT/Copilot (Bing Chat)/Bard
- Learn how to write prompts or get prompts
- [ChatGPT Translate: How to Write Your Translation Prompt \(instantaiprompt.com\)](https://instantaiprompt.com)
- [2304.02182.pdf \(arxiv.org\)](https://arxiv.org/abs/2304.02182)
- Try out novel tools, e. g. <https://www.deepl.com/write>

DeepL: KI-Schreibassistent

The screenshot displays the DeepL website interface. At the top, the DeepL logo and a 'Write' button are visible. Below the header, a grey banner reads 'Bessere Texte im Handumdrehen' (Better texts in a flash) with the subtitle 'Schreiben Sie klar, präzise und fehlerfrei' (Write clearly, precisely and error-free). The main area shows a language selection dropdown on the left with 'Deutsch' selected. The central text area contains the German text: 'oder einfügen, um Vorschläge zu sehen'. To the right, a panel shows the translated text in British English: 'The research institution will send the encrypted audio file of the focus group to a transcriber via a secure online portal. This person will transcribe the interviews. A pseudonym will be used rather than your real name on the transcripts. Once the transcriber had returned transcripts to the research institution, they will delete their audio files.' The translated text is highlighted with a blue background. Below the text, there are icons for audio playback, undo, and redo. On the far right, there are icons for thumbs up, thumbs down, and a share icon.

DeepL Write

Bessere Texte im Handumdrehen
Schreiben Sie klar, präzise und fehlerfrei

Deutsch ^

Englisch (britisch)
Englisch (amerikanisch)
Deutsch ✓

oder einfügen, um Vorschläge zu sehen

Vort, um Synonyme zu erhalten oder einen Satz

Englisch (britisch) v

The research institution will send the encrypted audio file of the focus group to a transcriber via a secure online portal. This person will transcribe the interviews. A pseudonym will be used rather than your real name on the transcripts. Once the transcriber had returned transcripts to the research institution, they will delete their audio files.

The research organisation sends the encrypted audio file of the focus group to a transcriber via a secure online portal. This person will transcribe the interviews. A pseudonym will be used instead of your real name on the transcripts. Once the transcriber has returned the transcripts to the research organisation, they will delete the audio files.

Speaker icon, Undo, Redo, Thumbs up, Thumbs down, Share icon

AI for translation (and other tasks) in novel environments

- Captioning software (CaptionHub, Sonix)
 - AI voice recognition (creating the transcript and captions)
 - Manual review of source captions
 - MT or AI translation of captions
 - Manual review of translation
- To do:
 - Explore these new environments
 - Explore possibilities for most productive way of working (import/export)

AI embedded in a CAT environment

- Direct access to ChatGPT from a CAT tool
 - Definition/Explanation
 - Translation
 - Alternatives/Rephrasing
 - Spell-check/Proof-reading
 - Etc.
- Options to overcome shortcomings of machine translation
 - Easier trainability
 - Alternatives
 - Built-in proofreader
 - Context/cohesion

To do:

- Explore AI access in traditional CAT tools
- Explore new options:
 - Additional Tools: CoTranslatorAI, Custom.MT
 - New CATs: Wordscope, Bureau Works

Example: Wordscope

The screenshot displays the Wordscope web interface. The top navigation bar includes the Wordscope logo, a dropdown menu with 'Bootcam...', and other options like 'Project', 'Share', 'Translation', 'Download', 'My Account', and 'Help'. The main workspace is divided into two columns: 'Source Text' and 'Insert or Edit your translation:'. The 'Source Text' column contains three entries: 'Ipsos-MRBI will invite you to take part in this study.', 'They will also invite you to be part of a focus group.', and 'They will discuss the study with you and ask for your verbal consent to take part.'. The 'Insert or Edit your translation:' column shows the German translation for the third entry: 'Sie werde die Studie mit Ihnen besprechen und um Ihre mündliche Einwilligung zur Teilnahme bitten.'. A 'ChatGPT' button is visible at the bottom of the workspace. On the right side, a sidebar titled 'ChatGPT for translators' is open, showing a chat window with a message about the verb 'werden' and a list of buttons for various functions like 'Clear', 'Definition', 'Explain', 'Translate', 'Rephrase', 'Alternatives', 'Spelling', 'Proofread', 'Back Translation', and 'Terminology'. The bottom status bar shows '3% Words: 207 To-do: 201 Chars: 98 / 98 [98] Bootcam clinical study.docx' and an 'Upgrade your Subscription' button.

WORDSCOPE Bootcam... Project Share Translation Download My Account Help EN JK

Research study:

Source Text	Translation
000006 Ipsos-MRBI will invite you to take part in this study.	Ipsos-MRBI wird Sie einladen, an dieser Studie teilzunehmen.
000007 They will also invite you to be part of a focus group.	Sie werden Sie auch zur Teilnahme an einer Fokusgruppe einladen.
000008 Source Text: They will discuss the study with you and ask for your verbal consent to take part.	Insert or Edit your translation: Sie werde die Studie mit Ihnen besprechen und um Ihre mündliche Einwilligung zur Teilnahme bitten.

ChatGPT for translators
Type a question and press ENTER or a few words and click one of the shortcut buttons.

ChatGPT Settings Terminology Help

and um Ihre mündliche Einwilligung zur Teilnahme bitten.

Spelling

Sie wird die Studie mit Ihnen besprechen und um Ihre mündliche Einwilligung zur Teilnahme bitten.

<< COPY

- The verb 'werden' (will) was incorrectly conjugated. In this sentence, it should be in the third person singular form, 'wird'. The correct sentence is 'Sie wird die Studie mit Ihnen besprechen und um Ihre mündliche Einwilligung zur Teilnahme bitten.'

Clear Definition Explain Translate Rephrase Alternatives Spelling Proofread Back Translation Terminology

Track Changes Close

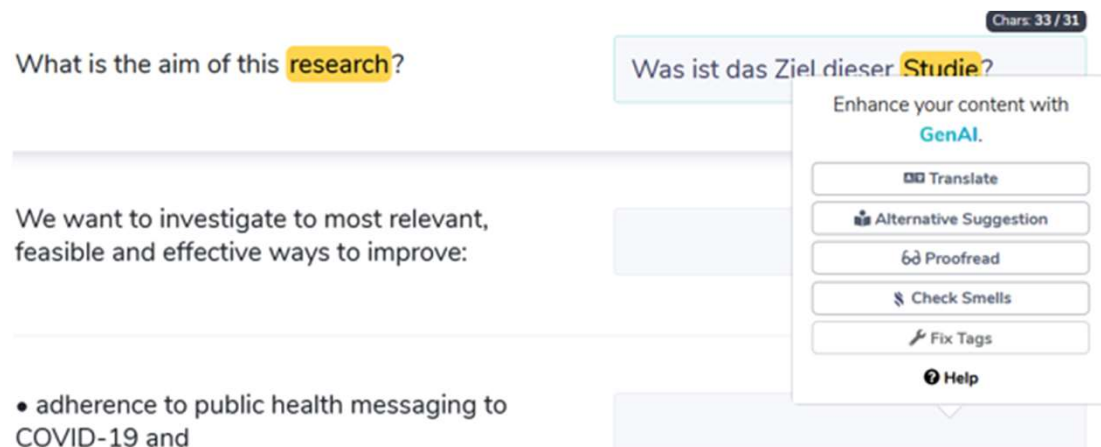
Upgrade your Subscription

Bureau Works

- Localization platform
- Online tool
- Features:
 - Translation management (TMs, glossaries, tasks)
 - Automated project management
 - **BWX AI**
 - Connectors/Integrations

Example: Bureau Works AI

- Combines glossaries, TMs, MT and ChatGPT
 - Translate
 - Alternative Suggestion
 - Proofread
 - Check Smells
 - Fix Tags



<https://support.bureauworks.com/hc/en-us/articles/18130511494797-All-about-our-AI-integration>

<https://www.loom.com/share/9eb2dbb2ea5746f7a4fdbfa2956a79c4> (Gabriel Fairman)

Example: Bureau Works

- Easier trainability
- Alternatives
- Built-in proofeditor
- Overcome cohesion problems

Use case: informed consent for research

What is the aim of this research?

We want to investigate the most relevant, feasible and effective ways to improve:

- adherence to public health messaging to COVID-19 and
- public health understanding and trust during the COVID-19 response.

What can I expect from taking part in this research study?

The field agency will invite you to take part in this study. They will also invite you to be part of a focus group. They will discuss the study with you and ask for your verbal consent to take part. If you agree to take part, they will then co-ordinate and establish a time that suits you to take part in the focus group.

You will be asked to complete a consent form and email it back directly to the research team.

This focus group will take approximately one hour and will involve a conversation with other people, using the teleconference software, Microsoft Teams. Interviews will be audio-recorded using Microsoft Voice recorder.

The focus group will be an informal conversation on your experiences and understanding of COVID-19, with topics including:

- how you and your family have coped during the pandemic,
- which sources you rely on and trust to gain information on COVID-19, and
- your views and opinions on responses to the pandemic.

You will receive €40 for your participation in the form of a pre-paid card.

If you wish to stop at any point during the focus group you may do so without giving an explanation.

- Legal requirements:
 - Client consent
 - Terms of Use of AI provider
 - Personal data
- Translational/textual requirements:
 - Mechanical issues:
 - Machine-readable, monolingual text
 - Cut-off sentences
 - Tags
 - Content issues:
 - Purpose (informative, not motivational, formal etc.)
 - Shift in audience
 - Cultural influence on topic/metaphors/way of presenting things
 - Language issues:
 - Language not neutral
 - Incorrect writing (complexity)
 - Form of address
 - Formatting conventions

Example: Bureau Works – easier trainability

1 What is the aim of this research?

Was ist das Ziel dieser Forschung?

Chars: 33 / 34

AI

⚡ BWX AI-Powered Actions

1 What is the aim of this research?

Was ist das Ziel dieser Studie?

Chars: 33 / 31

AI

⚡ BWX AI-Powered Actions

5 What can I expect from taking part in this research study?

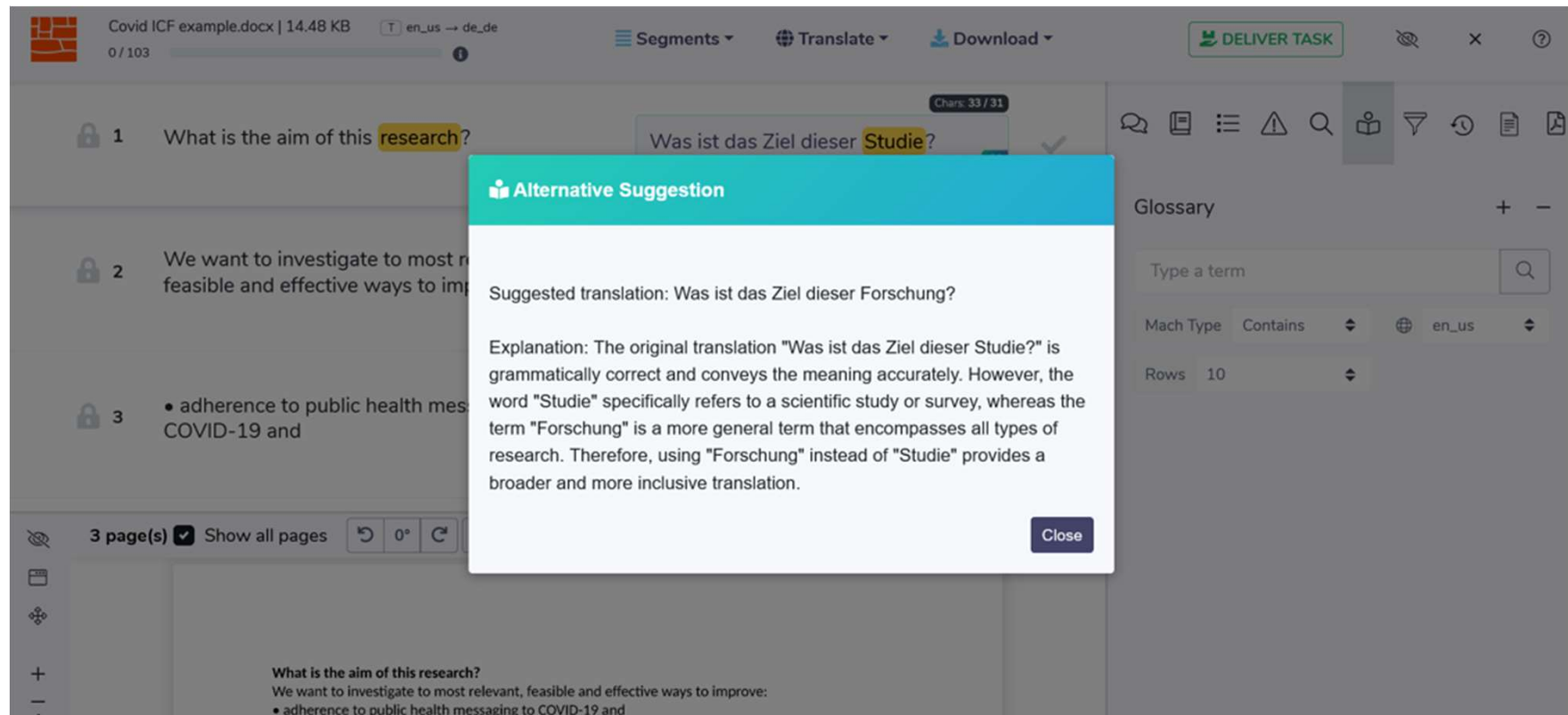
Was kann ich erwarten, wenn ich an dieser Studie teilnehme?

Chars: 58 / 59

AI

⚡ BWX AI-Powered Actions

Example: Bureau Works – alternative suggestions



Example: Bureau Works – built-in proofreader



6

The field agency will invite you to take part in this study.

Chars: 60 / 64

Die Feldagentur wir Sie einladen, an dieser Studie teilzunehmen.

AI

BWX AI-Powered Actions

AB Spell Checking Die **Feldagentur** wir Sie einladen, an dieser Studie teilzunehmen.

Translation Smells ⓘ

Grammatikfehler

Im deutschen Satz fehlt das Verb 'wird'. Der korrekte Satz sollte lauten: 'Die Feldagentur wird Sie einladen, an dieser Studie teilzunehmen.'

Example: Bureau Works – cohesion problems

21

The field agency will store your name and contact details in an encrypted file on a password-protected

Chars: 102 / 115

Die Feldagentur wird Ihren Namen und Ihre Kontaktdaten in einer verschlüsselten Datei auf einem passwortgeschützten

AI

BWX AI-Powered Actions

Spell Checking Die Feldagentur wird Ihren Namen und Ihre Kontaktdaten in einer verschlüsselten Datei auf einem passwortgeschützten

Translation Smells

Unvollständige Übersetzung Der deutsche Satz ist unvollständig. Es fehlt das Wort 'speichern' und der Kontext, auf was die Daten gespeichert werden (z.B. 'Computer', 'Server', 'Laufwerk'). Der vollständige Satz sollte lauten: 'Die Feldagentur wird Ihren Namen und Ihre Kontaktdaten in einer verschlüsselten Datei auf einem passwortgeschützten [Fehlendes Wort] speichern.'

22

computer for the purposes of co-ordinating the focus group interviews.

Chars: 70 / 67

-Computer speichern, um die Fokusgruppeninterviews zu koordinieren.

AI

BWX AI-Powered Actions

Spell Checking -Computer speichern, um die Fokusgruppeninterviews zu koordinieren.

Translation Smells

Fehlende Wörter Im deutschen Satz fehlt das Wort 'zum Zwecke der' oder 'für die Zwecke der', was im englischen Satz als 'for the purposes of' übersetzt wird.

Falsche Übersetzung Das Wort 'speichern' ist in der deutschen Übersetzung hinzugefügt worden, obwohl es im englischen Originalsatz nicht vorhanden ist. Es sollte entfernt werden.

Use case: challenges



1

Participants had in-person or phone call visits with the doctors once a week.

Die Teilnehmer hatten einmal pro Woche persönliche oder telefonische Besuche bei den Ärzten.

Chars: 77 / 92

AI

🔮 BWX AI-Powered Actions ▼

Alternative Suggestion

Suggested translation: Die Teilnehmer hatten einmal pro Woche persönliche oder telefonische Arztbesuche.

Explanation: In the original translation, the word "Besuche" was used, which means "visits." However, in this context, it is more appropriate to use the word "Arztbesuche," which specifically refers to visits to the doctors. This helps to clarify that the visits were related to medical appointments.

Close



Use case: challenges



1

Participants had in-person or phone call visits with the doctors once a week.

Chars: 77 / 95

Die Teilnehmer führten einmal pro Woche persönliche oder telefonische Gespräche mit den Ärzten.

AI

Translation Smells BWX AI-Powered Actions

Note: Translation Smells can show the presence of potential errors, but not ensure their absence.

Use case: challenges

	2	We call ourselves Birds!	Wir nennen uns selbst Birds!	AI	✓	
<input type="checkbox"/>		3	We work fast, grow fast, build fast and focus on impact.	Wir arbeiten schnell, wachsen schnell, bauen schnell und konzentrieren uns auf Wirkung.	AI	✓
	4	We're go-getters, industry leaders and roll-up-your-sleeves-and-make-it-happen kind of people.	Wir sind Macher, Branchenführer und packen an und setzen es um.	AI	✓	

Use case: challenges

 2 We call ourselves Birds!

Wir sind die Birds!

AI

 Translation Smells 

 Übersetzungsfehler

Die englische Phrase 'We call ourselves Birds!' wurde als 'Wir sind die Birds!' übersetzt. Eine genauere Übersetzung wäre 'Wir nennen uns Birds!'



☐  2 We call ourselves Birds!


Chars: 24 / 26

Hallo, wir sind die Birds!

AI

 Translation Smells

 BWX AI-Powered Actions 

 Hinzufügung Von Wörtern

Das Wort 'Hallo' wurde in der deutschen Übersetzung hinzugefügt, obwohl es im englischen Originaltext nicht vorhanden ist.



Use case: challenges



3

We work fast, grow fast, build fast and focus on impact.

Chars: 56 / 97

Unsere Arbeit, unser Wachstum und unsere Projekte haben Tempo und für uns zählen die Ergebnisse.

AI



Translation Smells

BWX AI-Powered Actions

Fehlende Übersetzung

Der Teil 'build fast' wurde nicht übersetzt. Eine mögliche Übersetzung könnte 'wir bauen schnell' sein.

Unnatürliche Übersetzung

Die Phrase 'haben Tempo' ist in diesem Kontext unnatürlich. Eine bessere Übersetzung könnte 'schnell voranschreiten' sein.



Value of the human translator – what makes the "human touch" so valuable?

- The translator understands the intention of your communication and will make sure that you can achieve it, even if this means changing the target audience or the medium. In short, your translator will make sure that you look good to your target audience...
- The translator acts as the first person in the target country to see the communication in question. A translator is almost like a mini focus group - they will notice what is awkward, inappropriate or missing.
- The translator will help you smooth out any elements of your communication that are culturally awkward or inappropriate. Often this happens without the client even realizing.
- The translator will proofread your text like no other proofreader. They need to really understand the text to be able to translate it, so anything unclear (or omitted) will be picked up - helping you to improve your communication in the source language.

Thank you!

And special thanks to Karen Leube, Jaqueline Jugenheimer and André Pires da Silva for their encouragement!

Further reading

- <https://arxiv.org/abs/2302.09210v1>
- <https://inten.to/blog/gpt-3-translation-capabilities/>
- <https://www.atanet.org/client-assistance/machine-translation-vs-human-translation/>

Back-up slide:

Full list of factors to consider for MT

- Legal requirements:
 - Client consent
 - Terms of Use of MT provider
 - Personal data
 - Know how and where your use of MT is recorded
- Technical requirements:
 - Selection of appropriate MT engine
 - Language pair/locale?
 - Option to train?
 - Fluency of output?
 - Integration into the tool of your choice
- Translational/textual requirements:
 - Mechanical issues:
 - Machine-readable, monolingual text
 - Cut-off sentences
 - Tags
 - Content issues:
 - Purpose (informative, not motivational, formal etc.)
 - Shift in audience
 - Cultural influence on topic/metaphors/way of presenting things
 - Language issues:
 - Language not neutral
 - Incorrect writing (complexity)
 - Form of address
 - Formatting conventions

Back-up slide:

Clarifications/terms

- **Artificial intelligence (AI)** is the intelligence of machines or software, as opposed to the intelligence of humans or other animals.
(https://en.wikipedia.org/wiki/Artificial_intelligence)
- **Generative artificial intelligence (generative AI, GAI, or GenAI)** is artificial intelligence capable of generating text, images, or other data using generative models, often in response to prompts.
(https://en.wikipedia.org/wiki/Generative_artificial_intelligence)

Back-up slide:

Clarifications/terms

- A **large language model (LLM)** is a language model notable for its ability to achieve general-purpose language generation. LLMs acquire these abilities by learning statistical relationships from text documents (...).
(https://en.wikipedia.org/wiki/Large_language_model)
- **Machine translation** is use of either rule-based or probabilistic (i.e. statistical and, most recently, neural network-based) machine learning approaches to translation of text or speech from one language to another, including the contextual, idiomatic and pragmatic nuances of both languages. (https://en.wikipedia.org/wiki/Machine_translation)

Back-up slide:

Shortcomings of Machine Translation

- Content
 - Mistranslations
 - read the text aloud
 - Convoluted wording
 - check the translation length (manually or with QA)
 - Added/missing information
 - check the translation length (manually or with QA)
- Mechanical errors
 - Glitches
 - QA
 - Tags
 - QA
 - Formatting conventions
 - regular expressions
- Stylistic errors
 - Repetitions
 - terminology functions, filtering
 - Sentence structure too literal
 - check the translation length (manually or with QA)
 - Inconsistencies
 - terminology functions, filtering
 - Terminology
 - terminology functions
 - Cohesion ("it")
 - filter for "It"
 - Form of address
 - filter for appropriate words ("Du")
- Tracking of use of MT: you need to know where this information is stored
 - know your CAT tool