

A LINGUISTIC QUANTUM LEAP

LLMs in Localization and Translation

MARINA PANTCHEVA | DECEMBER 8, 2023



Director Linguistic
AI Services at RWS

- PhD on Theoretical Linguistics
- 2014: transition to the **localization** industry
- 2015-2021: creation of a **translation platform for translators**.
- 2021-2023: leading a group that develops **solutions for crowd localization**
- 2023: in charge of Linguistic AI research and implementation of **Linguistic AI Services**

LOCALIZATION MEANS

**TRANSLATING A PRODUCT PLUS
ADAPTING THE TRANSLATION TO
A SPECIFIC REGION.**

LOCALIZATION IS MORE THAN JUST TRANSLATION

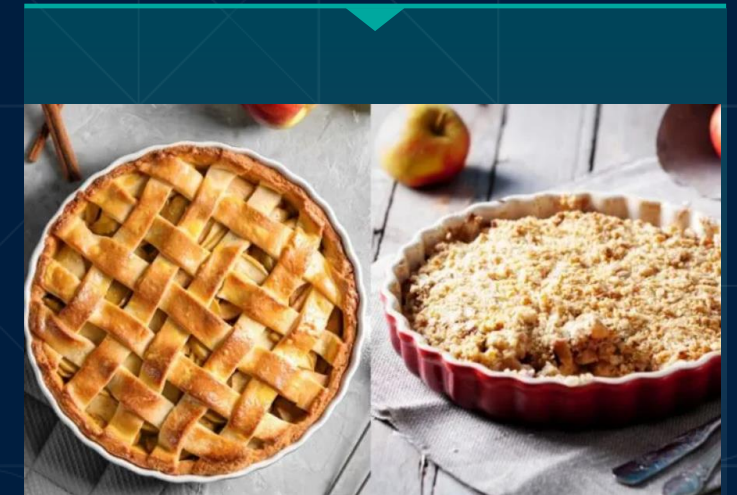
CONVERSION OF TEXT FROM ONE LANGUAGE TO ANOTHER

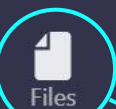
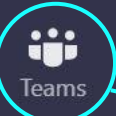
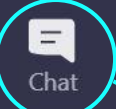


ADAPTATION TO THE MARKET AND CULTURE

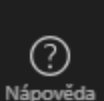
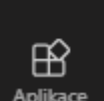
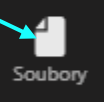
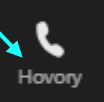
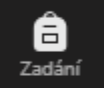
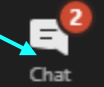
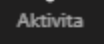


TRANSFORMATION OF NON-TEXTUAL COMPONENTS



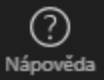


Váš stav je nastaven na Nerušit. Budete dostávat upozornění pouze na naléhavé zprávy a zprávy od prioritních kontaktů. [Změnit nastavení](#)



- Zobrazit nastavení zařízení
- Zobrazit poznámky
- Zobrazit podrobnosti o schůzce
- Zahájit režim celé obrazovky
- Zobrazit efekty pozadí
- Zapnout živé titulky (verze Preview, jenom americká angličtina)
- Klávesnice
- Spustit nahrávání
- Ukončit schůzku
- Vypnout přechozí video

02:03 [Mute] [Microphone] [Screen Share] [More] [Hand] [Chat] [Participants] [End Call]



TERMINOLOGY DATABASES AND STYLE GUIDES

Terminology: a set of specialized terms used within a specific industry domain or product and their approved translations.

Style guide: a set of rules and guidelines that dictate the preferred translation style for a particular brand or product.

TRANSLATION MEMORIES

A database of source text and its corresponding translation.

It contains all **segments** (sentences, paragraphs or sentence-like units) that **have been translated before**.

It helps translators **translate faster** and **more consistently**.

COMPUTER ASSISTED TRANSLATION TOOLS

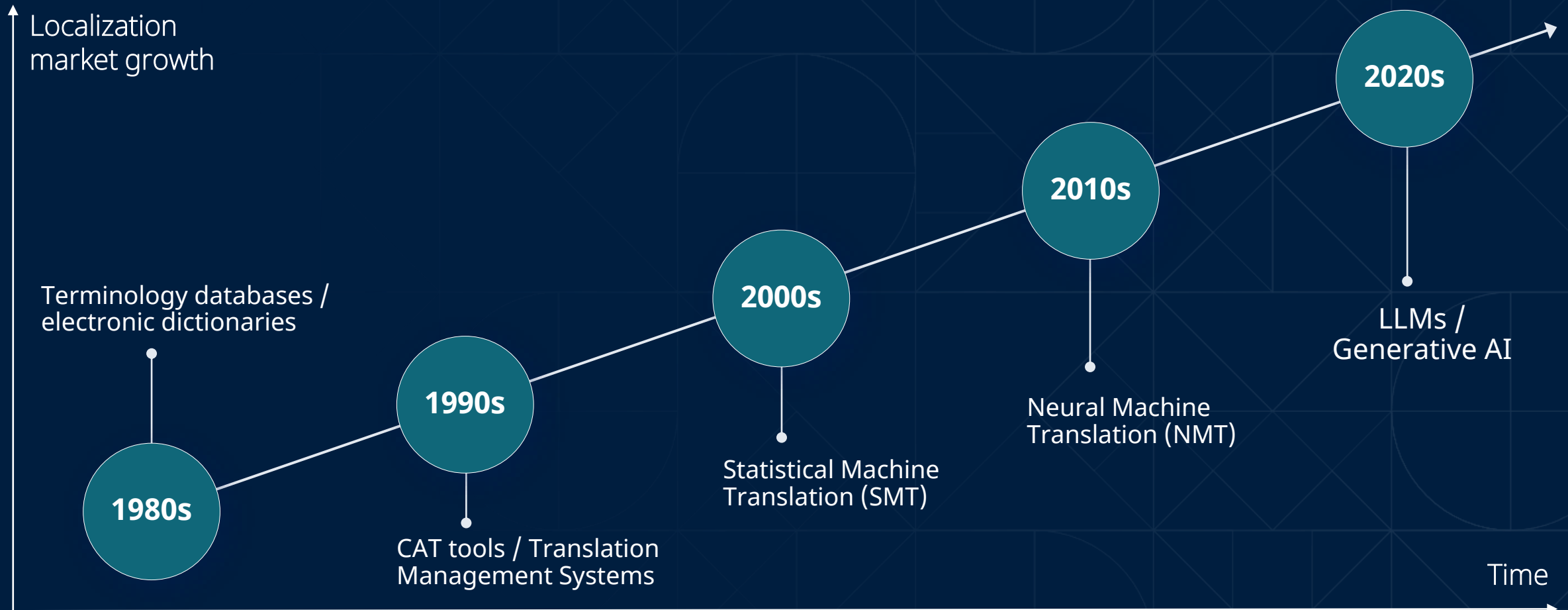
An **application** designed to help human translators in the process of translating text from one language to another.

CAT tools provide features such as translation memory, terminology management, and alignment of source and target texts.

MACHINE TRANSLATION

The automated process of translating text from one language to another using computer algorithms.

LOCALIZATION INDUSTRY EVOLUTION



ARTIFICIAL INTELLIGENCE

ARTIFICIAL INTELLIGENCE

A program that can sense, reason, act, and adapt

DEEP LEARNING

Subset of machine learning in which multilayered neural networks learn from vast amounts of data

LARGE LANGUAGE MODELS

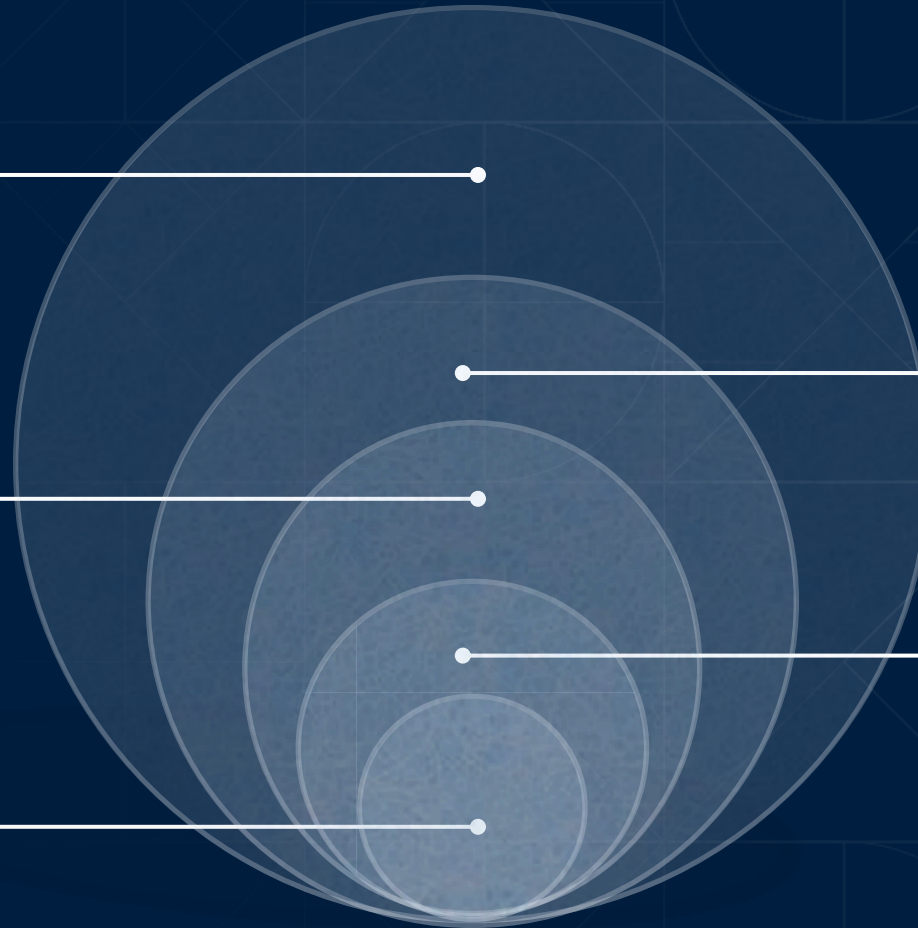
AI that is trained on vast amounts of **text** to interpret and generate human-like textual output

MACHINE LEARNING

Algorithms whose performance improves as they are exposed to more data over time

GENERATIVE AI

AI that learns from data about existing artefacts and then uses the knowledge to generate new artefacts



WHAT CAN WE USE LLMS FOR?

There are *numerous ways* in which LLMs **can HELP** increase the speed, quality and efficiency of the localization process.

USING LLMs FOR TRANSLATION

TRADOS OPEN AI TRANSLATOR

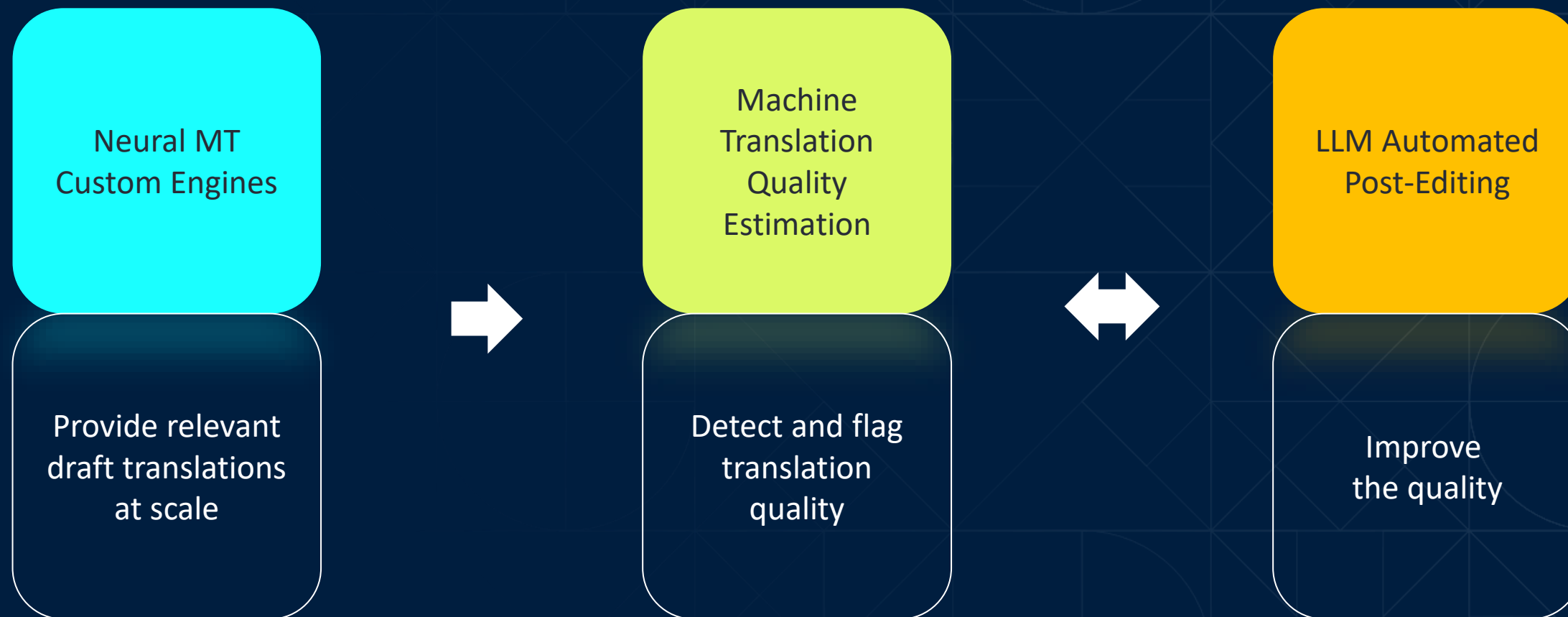
The screenshot displays the Trados OpenAI Translator interface. On the left, the source text in Hebrew is shown in a window titled 'openAI_he.txt'. On the right, the target text in English is shown in a window titled 'openAI_he.txt'. The comparison window on the right shows the original English text and the translated Hebrew text, with red circles and arrows highlighting specific differences and features.

Segment	Source (Hebrew)	Target (English)
1	בעבר, OpenAI פותחה כדי לעזור לבני אדם להשיג דברים גדולים באמצעות בינה מלאכותית.	Once upon a time, OpenAI was developed to help humans achieve great things using artificial intelligence.
2	עם זאת, הצוות מאחורי OpenAI הבין במהירות כי הטכנולוגיה שלהם יכולה לעשות יותר מאשר רק crunch נתונים ולפתור בעיות מורכבות.	However, the team behind OpenAI quickly realized that their technology could do more than just crunch data and solve complex problems.
3	יום אחד, הם החליטו לחקור כיצד OpenAI יכולה להביא יצורים שונים יחד בהרמוניה.	One day, they decided to explore how OpenAI could bring different creatures together in harmony.
4	הם התחילו עם חתולים, כלבים וציפורים - שלושה בעלי חיים הידועים לשמצה בכך שאינם מסתדרים.	They started with cats, dogs, and birds - three animals that are notoriously known for not getting along.
5	באמצעות יכולות עיבוד השפה הטבעית של OpenAI, פיתח הצוות מערכת שיכולה לתרגם בין שפות של חתולים, כלבים וציפורים.	Using OpenAI's natural language processing capabilities, the team developed a system that could translate between the languages of cats, dogs, and birds.

The comparison window on the right shows the original English text and the translated Hebrew text, with red circles and arrows highlighting specific differences and features:

- 1: Prompt: Multiple transl... Search from Source
- 2: Translations: However, the team behind OpenAI quickly realized that their technology could do more than just crunch data and solve complex problems.
- 3: HoweverYet, the team behind-OpenAI quicklyteam realizedswiftly comprehended that their technology couldwas de capable of more than just erunchprocessing data and solveresolving complexintricate problemsissues.
- 4: Model: gpt-3.5-turbo
- 5: Show differences

THE POWER OF MACHINE TRANSLATION AND LLM IN ONE LANGUAGE WEAVER | EVOLVE



LINGUISTIC AI SERVICE: TERM CHANGE IMPLEMENTATION

THE HUMAN APPROACH TO TERMINOLOGY CHANGES

There's a change of the official CZ translation for the term "view" (v) from **zobrazit** to **ukázat**.

How do we implement this change in the Translation Memory?

Go through the English-Czech pairs in the Translation Memory and replace any occurrence of the outdated term **zobrazit** by **ukázat**.

While this works for (most) humans, it would make a very bad instruction for an LLM.

TERM CHANGE IMPLEMENTATION PROMPTING

THINKING STEP-BY-STEP

STEP 1

DETECT IF THE SOURCE TERM IS PRESENT



STEP 2

IDENTIFY IF THE TRANSLATION USES THE APPROVED TARGET TERM



STEP 3

REPLACE THE OLD TARGET TERM BY THE NEW ONE

Check if the concepts match

Check for derivations and inflections

Check the part of speech

Check for partial match
(Is the term part of a larger phrase?)

Check if the target term is present

Consider the changes that the term can undergo capitalization, morphosyntactic changes, compounding

Replace the outdated translation by the new term adjusting the latter to the syntactic structure it is plugged into.

THE PROMPT FOR STEP 2

SMART TARGET TERM CHECK [STEP 2]

ROLE You are a professional {{Language}} reviewer. Your task is to evaluate if a {{Language}} sentence contains the approved terminology translations.

SOURCE-TARGET PAIR

Your task is to evaluate if a {{Language}} sentence uses the approved translation {{Target_Term}} for the source term {{Source_Term}}.

ADDITIONAL INSTRUCTIONS

Do not match the word {{Target_Term}} one-to-one but keep in mind that the word may change in the translation because of case endings, number agreement, gender agreement, verbal tenses, different capitalization.

TASK

You must evaluate the following pair:
English sentence: {{Source_Segment}}
Translation: {{Target_Segment}}

DEFINE OUTPUT

If the translation contains the approved term {{Target_Term}} in some morphological form, then provide as output "No error".

If the translation does not use the approved term {{Target_Term}} at all, then provide as output "Error".

REDUCE NOISE

Do NOT return any other response "No error" or "Error".

EXAMPLE: SMART TARGET TERM CHECK [PROMPT 2]

Use LLMs to detect if the translation contains the approved target term regardless of the target term's morphosyntactic form.

	RUSSIAN	FINNISH!
GPT 3.5	75–90%	80–90%
GPT 4	100%	100%

Run on a sample that contains 100% relevant source term hits.

PROMPT ENGINEERING 101

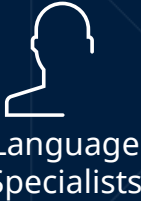
The art of designing linguistic prompts

Linguistic prompts are **not universal**.
They must be **custom made** for every use
case, client, text type and language.

WHAT IS A PROMPT?

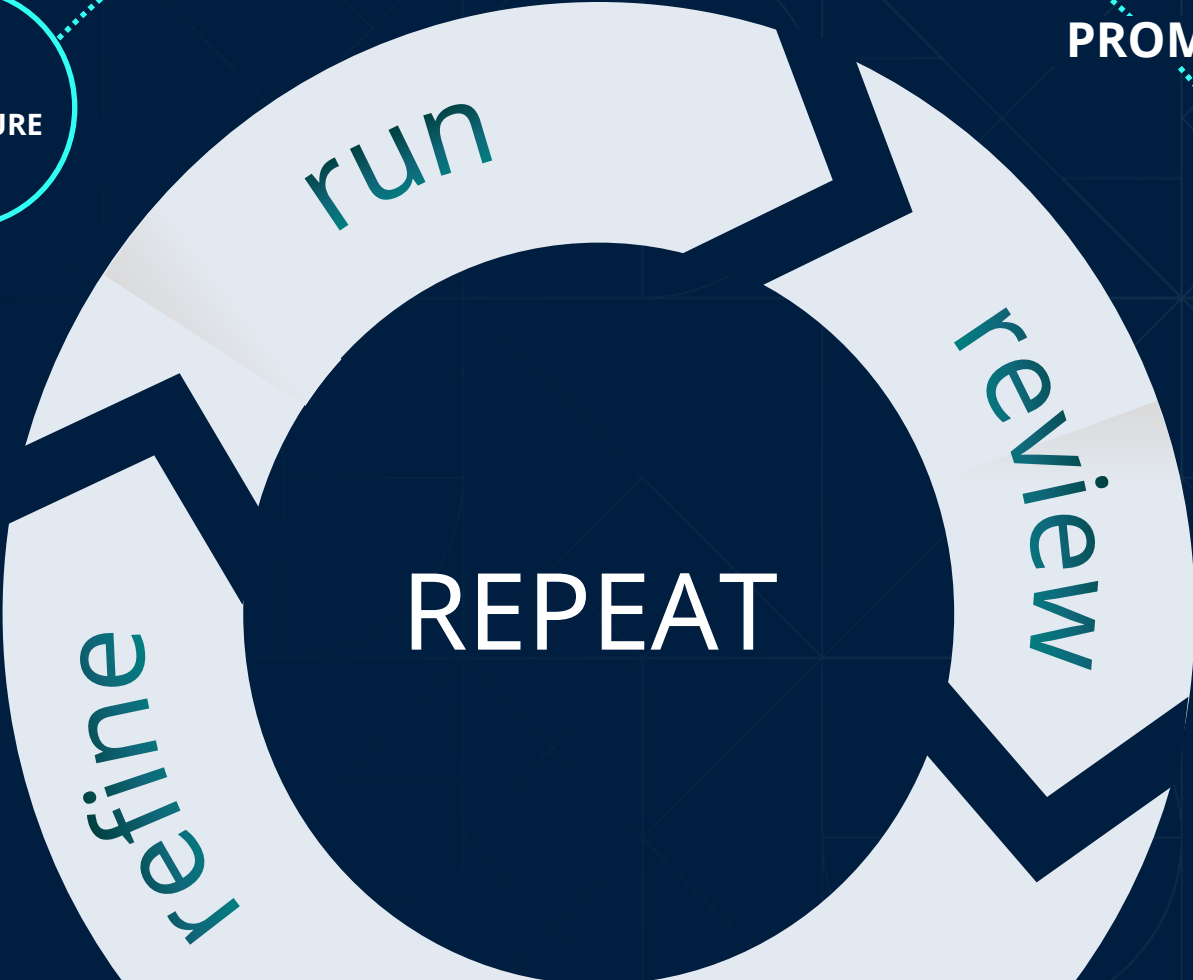
An instruction that initiates a response or an action by an LLM.

PROMPT CREATION



MASTER PROMPT

PROMPT ELEMENTS



Creating a prompt involves iterative cycles of response evaluation and prompt refinement.



Linguists are instrumental in adjusting the prompt to the language specificities of their language.

THE CONTRIBUTION OF LANGUAGE SPECIALISTS

CUSTOMIZED LANGUAGE PROMPTS

LANGUAGE SPECIALISTS

MASTER PROMPT

Defining the prompting strategy, elements and structure

German prompt

Spanish prompt

Polish prompt

Japanese prompt

Bulgarian prompt



EXAMPLE: MASTER PROMPT

DETECTING TONE OF VOICE

ROLE

You are a professional {{Language}} linguist. Your task is to evaluate if a {{Language}} sentence addresses the customer in a formal or informal way

INTRO

Take a breath and do a step-by-step analysis.

DETECT ADDRESS

Start by examining the sentence to determine if it addresses the customer directly. You can determine this by detecting second person pronouns or imperative verb forms. If the sentence contains second person pronouns {{examples}}, then it addresses the customer directly.

DETERMINE TONE OF VOICE

As a next step analyze the form of verbs and pronouns to determine whether the customer is addressed formally or informally. Consider personal pronouns, possessive pronouns, verb agreement, and the specific imperative form of the verb

SHOTS

Here come a few examples:

The sentence {{Example 1}} is "Informal". The sentence {{Example 2}} is "Formal".

TASK

Apply the instructions above to the given sentence: "{{SEGMENT}}".

DEFINE OUTPUT

If the sentence doesn't address a customer directly but is a generic statement, respond with "NA".

If the sentence addresses the customer formally, reply with "Formal".

If it addresses the customer informally, reply with "Informal".

REDUCE NOISE

Do NOT provide any other responses than JUST "Formal", "Informal", or "NA".

CHALLENGES

LANGUAGE ADJUSTMENTS

GERMAN

Pronominal and verb form ambiguity between **they** (3p.pl) and **You** (2p.sg/pl formal)

SPANISH

Verb form ambiguity between **he/she** (3p.sg) and **You** (2p.sg formal)

DUTCH

Mixed address within a single segment!

ITALIAN

Ambiguity between **she** (3p.fem.sg) and **You** (2p.sg/pl formal);
Usage of impersonal constructions in target (e.g. "You can ..." -> "It is possible to...")

THE GERMAN PROMPT

DETECTING TONE OF VOICE

ROLE

You are a professional **German** linguist. Your task is to evaluate if a **German** sentence addresses the customer in a formal or informal way

INTRO

Take a breath and do a step-by-step analysis.

DETECT ADDRESS

Start by examining the sentence to determine if it addresses the customer directly. You can determine this by detecting second person pronouns or imperative verb forms. If the sentence contains second person pronouns (such as **du, dein, dir, dich, Sie, Ihr, Ihren**) then it addresses the customer directly.

DETERMINE TONE OF VOICE

As a next step analyze the form of verbs and pronouns to determine whether the customer is addressed formally or informally. Consider personal pronouns, possessive pronouns, verb agreement, and the specific imperative form of the verb. **Keep in mind that when the pronoun \"Sie\" starts with a capital letter in the middle of a sentence, this is an indication that the customer address is formal.**

SHOTS

Here come a few examples:

The sentence "**Lade dein Video hoch**" is "Informal". The sentence "**Laden Sie Ihr Video hoch**" is "Formal".

TASK

Apply the instructions above to the given sentence: "**{{SEGMENT}}**".

DEFINE OUTPUT

If the sentence doesn't address any customer directly but is a generic statement, respond with "NA".

If the sentence addresses the customer formally, reply with "Formal".

If it addresses the customer informally, reply with "Informal".

REDUCE NOISE

Do NOT provide any other responses than JUST "Formal", "Informal", or "NA".

THE DUTCH PROMPT

DETECTING TONE OF VOICE

ROLE

You are a professional **Dutch** linguist. Your task is to evaluate if a **Dutch** sentence addresses the customer in a formal or informal way

INTRO

Take a breath and do a step-by-step analysis.

DETECT ADDRESS

Start by examining the sentence to determine if it addresses the reader directly. You can determine this by detecting second person pronouns or imperative verb forms. If the sentence contains second person pronouns (such as **jij, je, jouw, u, uw**) then it addresses the customer directly.

DETERMINE TONE OF VOICE

As a next step analyze the form of verbs and pronouns to determine whether the customer is addressed formally or informally. Consider personal pronouns, possessive pronouns, verb agreement, and the specific imperative form of the verb.

SHOTS

Here come a few examples:

The sentence "**Sleep je media naar de tijdlijn onder je video**" is "Informal". "**Leert u hoe je voice-overs maak**" is "Mixed" because it uses both "**u**" and "**je**" in the same sentence. The sentence "**Met de functie voor automatische ondertiteling kunt u ondertitels in één taal per video genereren.**" is "Formal".

TASK

Apply the instructions above to the given sentence: "**{{SEGMENT}}**".

DEFINE OUTPUT

If the sentence doesn't address any customer directly but is a generic statement, respond with "NA".

If the sentence addresses the customer formally, reply with "Formal".

If it addresses the customer informally, reply with "Informal".

If you detect both formal and informal language, then reply "Mixed"

REDUCE NOISE

Do NOT provide any other responses than JUST "Formal", "Informal", "**Mixed**" or "NA".

THE SPANISH PROMPT

DETECTING TONE OF VOICE

- ROLE** You are a professional **Spanish** linguist. Your task is to evaluate if a **Spanish** sentence addresses the customer in a formal or informal way
- INTRO** [Take a breath and do a step-by-step analysis.
- DETECT ADDRESS** [Start by examining the sentence to determine if it addresses the customer directly. You can determine this by detecting second person pronouns or imperative verb forms. If the sentence contains second person pronouns (**tu, tus, te, usted, su, se**) then it addresses the customer directly.
- DETERMINE TONE OF VOICE** [As a next step analyze the form of verbs and pronouns to determine whether the customer is addressed formally or informally. Consider personal pronouns, possessive pronouns, verb agreement, and the specific imperative form of the verb.
- SHOTS** [Here come a few examples:
The sentence "**Tu cuenta está programada para su eliminación.**" is "Informal". The sentence "**No se pudo enviar un correo electrónico dirigido a usted.**" is "Formal".
- TASK** [Apply the instructions above to the given sentence: "**{{SEGMENT}}**".
- DEFINE OUTPUT** [If the sentence doesn't address any customer directly but is a generic statement, respond with "NA".
If the sentence addresses the customer formally, reply with "Formal".
If it addresses the customer informally, reply with "Informal".
- REDUCE NOISE** [Do NOT provide any other responses than JUST "Formal", "Informal", or "NA".

AN EVEN BETTER SPANISH PROMPT

DETECTING TONE OF VOICE

ROLE	Eres un lingüista profesional de español. Tu tarea es evaluar si una oración en español se dirige al lector de manera formal o informal.
INTRO	Tómate un momento y realiza un análisis detallado.
DETECT ADDRESS	Comienza examinando la oración para determinar si se dirige directamente al lector. Puedes determinarlo detectando pronombres de segunda persona o formas verbales en imperativo. Si la oración contiene pronombres de segunda persona (tú, tus, te, usted, su, se), entonces se dirige directamente al lector.
DETERMINE TONE OF VOICE	Como siguiente paso, analiza la estructura de los verbos y pronombres para determinar si la oración se dirige al lector de manera formal o informal. Considera pronombres personales, posesivos, concordancia verbal y la forma del verbo.
SHOTS	Aquí tienes algunos ejemplos: La frase "Tu cuenta está programada para su eliminación." es "Informal". La frase "No se pudo enviar un correo electrónico dirigido a usted." es "Formal".
TASK	Aplica las instrucciones anteriores a la oración: "{{SEGMENT}}".
DEFINE OUTPUT	Si la oración no se dirige directamente al lector y es una afirmación genérica, responde con "NA". Si la oración se dirige al lector de manera formal, responde con "Formal". Si se dirige al lector de manera informal, responde con "Informal".
REDUCE NOISE	NO proporciones ninguna otra respuesta que no sea "Formal", "Informal" o "NA".

Engine: DTX_NSwiss GPT-4

Sampling: Temperature

Frequency Penalty: 1

Presence Penalty: 0

Test round: [redacted]_mixedTM_formal_informal_check_2211_02

Evaluation method: Boolean

Test case id	Input	Completion	Is model response correct?	Tokens	Cost
1	[redacted] se preocupa por su privacidad. Para ayudarnos a protegerla, algunas características solo pueden usarlas ciertas pers Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	425	0,01
10	Se está comprobando que es usted Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	319	0,01
11	Los delegados pueden hacer y recibir llamadas en tu nombre. Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	323	0,01
12	Este chat es de uso exclusivo para usted. Úselo para borraros, enviarse archivos o conocer las características del chat un poc Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	341	0,01
13	Cuando bloquea a alguien, esa persona no podrá ver su estado y usted no recibirá ninguna comunicación de [redacted] por parte Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	356	0,01
14	Esta persona es de otra organización. Cuando aceptes, parte de la información de tu cuenta se compartirá con esta organizaci Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	361	0,01
15	El idioma de los subtítulos se ha establecido en %1\$s. Esto solo le afecta a usted. Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	335	0,01
16	Solo tenemos dos preguntas para usted. Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	318	0,01
17	%1\$s no usa [redacted] con una cuenta administrada por la organización Spanish	NA	<input checked="" type="radio"/> Yes <input type="radio"/> No	325	0,01
18	Sea amable y respetuoso con sus compañeros de equipo. No sea grosero ni cruel. Sea usted mismo y no publique nada que ir Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	351	0,01
19	%1\$s ha aceptado tu invitación a %2\$s Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	324	0,01
2	¿Estás seguro de que deseas cerrar la sesión? Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	323	0,01
20	Hemos seleccionado previamente el canal %1\$s para usted. \nElija un canal diferente aquí. Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	335	0,01
21	%s ha aceptado esta solicitud Spanish	NA	<input checked="" type="radio"/> Yes <input type="radio"/> No	317	0,01
22	Agregar tu propia categoría Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	316	0,01
23	Ya has solicitado unirse a esta comunidad. Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	322	0,01
24	Ya ha solicitado unirse a esta comunidad.	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	321	0,01
25	Cualquier persona puede unirse a la comunidad Spanish	NA	<input checked="" type="radio"/> Yes <input type="radio"/> No	320	0,01
26	Nos gustaría que compartieses datos de diagnóstico opcionales sobre cómo usas [redacted] Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	331	0,01
27	Cuando aceptes, parte de la información de tu cuenta se compartirá con esta organización. <a href="https://[redacted]" Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	354	0,01
28	[redacted] es un servicio de colaboración empresarial. Usa solo la información de la cuenta que proporcionó tu administrador de T Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	397	0,01
29	Únete a el primer equipo Spanish	Informal	<input checked="" type="radio"/> Yes <input type="radio"/> No	318	0,01
3	Si bloquea a alguien, esa persona no podrá ver su estado y usted no recibirá ninguna comunicación de Teams o Skype por par Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	358	0,01
30	La comunidad que está buscando puede ser privada. Obtenga un vínculo para unirse o crea la propia comunidad. Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	337	0,01
4	Solo tenemos dos preguntas para usted. Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	318	0,01
5	Esto eliminará permanentemente el chat para usted, pero no para %s. Spanish	Formal	<input checked="" type="radio"/> Yes <input type="radio"/> No	327	0,01

Pronoun drop and ambiguity

No need for overt pronoun "tu"

Correct parsing of reflexive

100% correct! Response correct Yes: 30 (100%), No: 0 (0%)

Temperature: 1
 Frequency penalty: 0
 Presence penalty: 0
 Total cost: Tokens 10135 (10079 prompt, 56 completion) \$ 0,31

DETECTING TONE OF VOICE

Following a change in style guide rules, employ an LLM to detect which TM segments use an outdated (formal) tone of voice.

	GERMAN	DUTCH	SPANISH	ITALIAN
GPT 3.5	70–90%	80–90%	80–90%	60-70%
GPT 4	100%	100%	100%	wip

AZURE OPENAI GPT

ERRORS SEEN IN TESTS



RESPONSE

Sometimes LLMs responses are excessively helpful. The response contains more information than is requested.



REPLICABILITY

The same prompt can elicit different responses at different times.



CONTEXT

The response may miss some features of the input and be incorrect in the context of the task.



LANGUAGE

Language-agnostic prompts can generate varying quality of response in different languages.

PROMPTING BEST PRACTICES



BE CLEAR

Write prompts in a simple and unambiguous language. Be explicit.



GIVE EXAMPLES

Use plentiful examples: TMs matches, terminology and any available resources for a few-shot prompting.



CLEAN THE DATA

Make sure the input dataset is prepared to optimally suit the use case.



BE NEGATIVE

To reduce noise and ensure clean output, clearly state what the model **MUST NOT** do.



THINK STEP BY STEP

Break down the task into discrete units and instruct the model to proceed step-by-step
(Chain-of-thought)

WHAT ELSE CAN WE USE LLMS FOR?

AI LQA

Evaluating the translation quality

SOURCE OPTIMIZATION

Improving the source text in terms of language and style

TERM EXTRACTION

Extracting key term candidates from a source text

VOICE CONVERSION

Converting formal language into informal

BEST TRANSLATION

Identifying the best translation for a single source segment out of multiple available translations in the translation memory

**VARIANCE
GENERATION**

Providing multiple translation variants for a given source

SMART TERM CHECK

Detecting terminology issues

**TERM CHANGE
IMPLEMENTATION**

Replacing an outdated term by a new one

HALLUCINATIONS

There are *numerous ways* in which LLMs **SURPRISE** us...

AZURE OPENAI GPT

HALLUCINATIONS



TECHNICALLY CORRECT BUT UNEXPECTED RESPONSE

Example: Providing Spanish variant-specific language (e.g. es-ARG, es-ES) out of the blue.

Mitigation: Think of all the contextually implicit information contained in the request and spell it out.



TOTAL NONSENSE

Example: Purported accuracy error in the translation

Mitigation: few-shot learning

The translation inaccurately translates 'Policy wizard' as 'Assistant de politique' instead of 'Assistant de politique'.



FAILURE TO FOLLOW INSTRUCTIONS

Example: Instead of TRUE/FALSE the LLM responds with YES/NO or even with a full explanatory paragraph.

Mitigation: be explicit in the definition of desired output. Provide negative commands.



PLAIN LIES

Example: Stating a fact that is obviously false.

Mitigation: Chain-of-thought or thread-of-thought prompting.

*Source: [Site created](#) > Target: [Web se vytvořil](#).
"The translation is an overtranslation. The word 'Web' is not necessary and can be omitted."*

RESOURCEDNESS

The amount of training data available for a given language

“**High-resource** languages are expected to show better **performance.**”

Performance across languages

Accuracy



Resourcedness of languages

Resourcedness





- ❑ Internships, workshops and training
- ❑ Trados technology for teaching

