LanguageAI™

Steps in the process:

Intake Automation

AI analyzing customer content

Pre-translation

AI optimizing source content

MT Customization + Model Fine-Tuning

AI-optimized data used for MT training and LLM fine-tuning

AI Translation

Custom MT Engine and AI-powered Translation Memory

Automated Post-Editing

Rules, AI Glossary Replacement, and AI stylistic choices
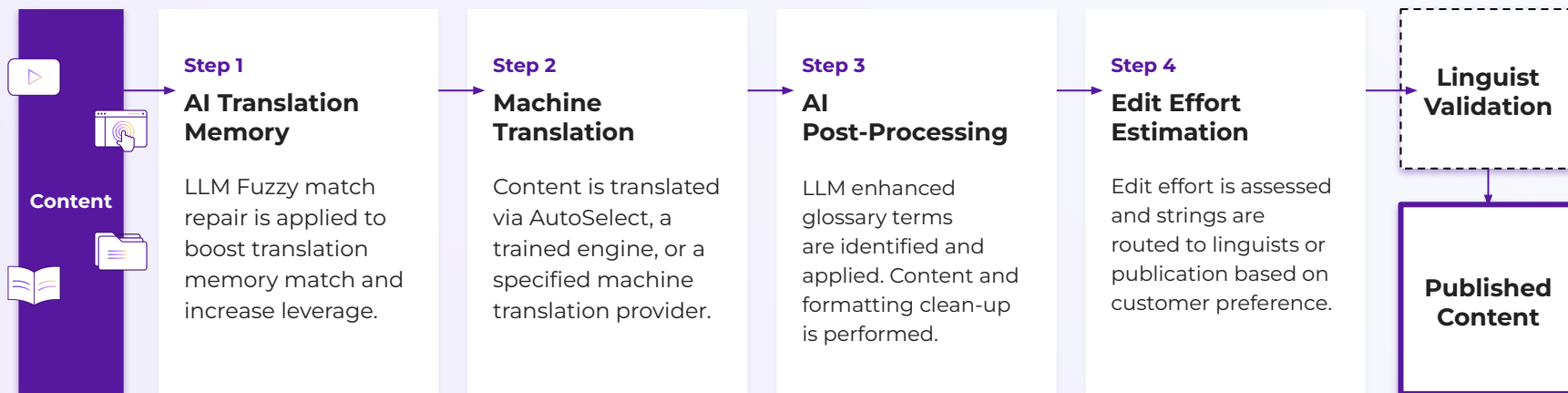
AI Workflow Decision

AI level of effort predictions and decisions

AI-powered Human Translation and Quality Evaluation

AI Translation Toolkit: How it works

Steps in the process:

Content

Step 1
AI Translation Memory

LLM Fuzzy match repair is applied to boost translation memory match and increase leverage.

Step 2
Machine Translation

Content is translated via AutoSelect, a trained engine, or a specified machine translation provider.

Step 3
AI Post-Processing

LLM enhanced glossary terms are identified and applied. Content and formatting clean-up is performed.

Step 4
Edit Effort Estimation

Edit effort is assessed and strings are routed to linguists or publication based on customer preference.

Linguist Validation

Published Content

AI Translation Toolkit: Early results

200%
Reduction in active linguist time

15%+
Decrease in number of edits

98+
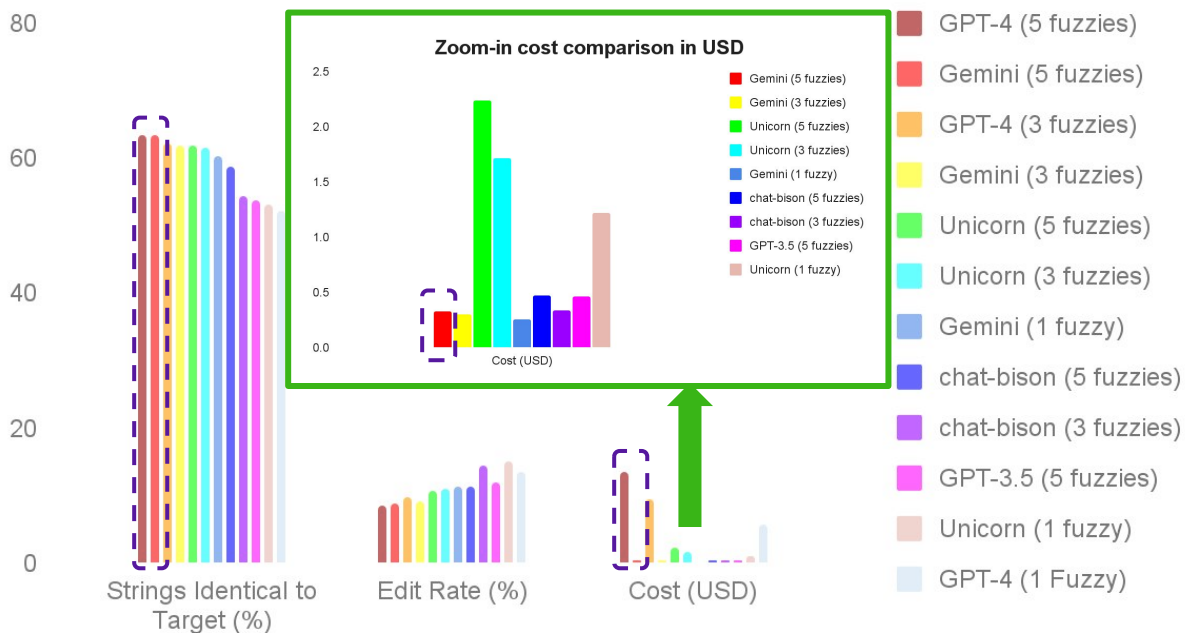**MQM score when paired with human validation**

Fuzzy Match Repair

Source Text	TM (Source)	TM (Target)
Are you sure you want to delete this **chat**?	Are you sure you want to delete this **photo**?	¿Seguro que quieres borrar **esta foto**?

LLMs

Suggested Translation

¿Seguro que quieres borrar **este chat**?
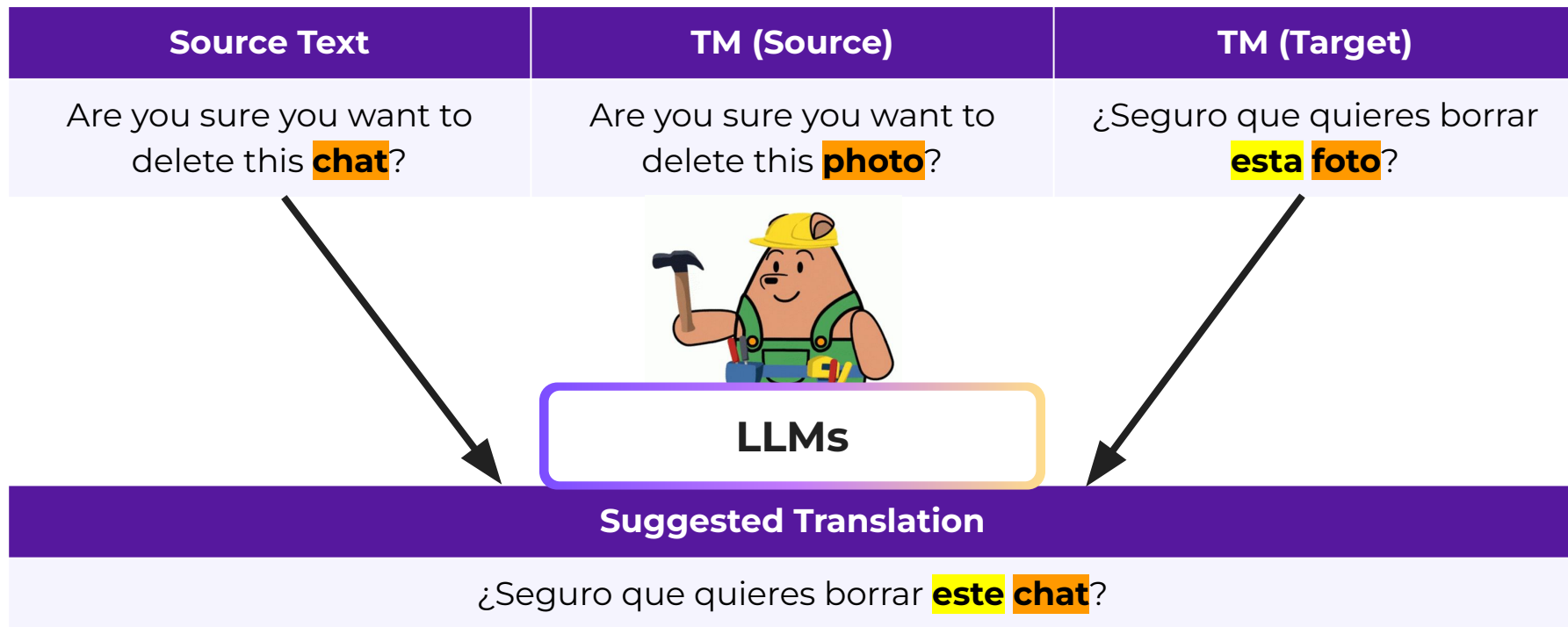
Fuzzy Match Repair LLM bake-off

Task: Fuzzy Match Repair

Providing the prompt with similar examples in the clients TM to yield the desired translation to the current source text.

Results:

Google's models yielded the best performance-cost tradeoff. Compared to a competitor's most competitive model, Google's Gemini yielded the same % strings identical to target translation while being 40 times cheaper in price!

Top performers for Fuzzy Match Repair (zh-CN)

- GPT-4 (5 fuzzies)
- Gemini (5 fuzzies)
- GPT-4 (3 fuzzies)
- Gemini (3 fuzzies)
- Unicorn (5 fuzzies)
- Unicorn (3 fuzzies)
- Gemini (1 fuzzy)
- chat-bison (5 fuzzies)
- chat-bison (3 fuzzies)
- GPT-3.5 (5 fuzzies)
- Unicorn (1 fuzzy)
- GPT-4 (1 Fuzzy)

SMARTLING

Fuzzy Match Repair: In-Prod Performance

% Strings with HTER = 0

Legend: pre-LLM-FMR, post-LLM-FMR

Categories: es, fr, de, zh, pt, it, ja, nl, ko, pl

Glossary insertion

Source Text	MT	Glossary
Schedule your first post	Programe su **primer** **post**	{post : publicación}

Traditional

LLM
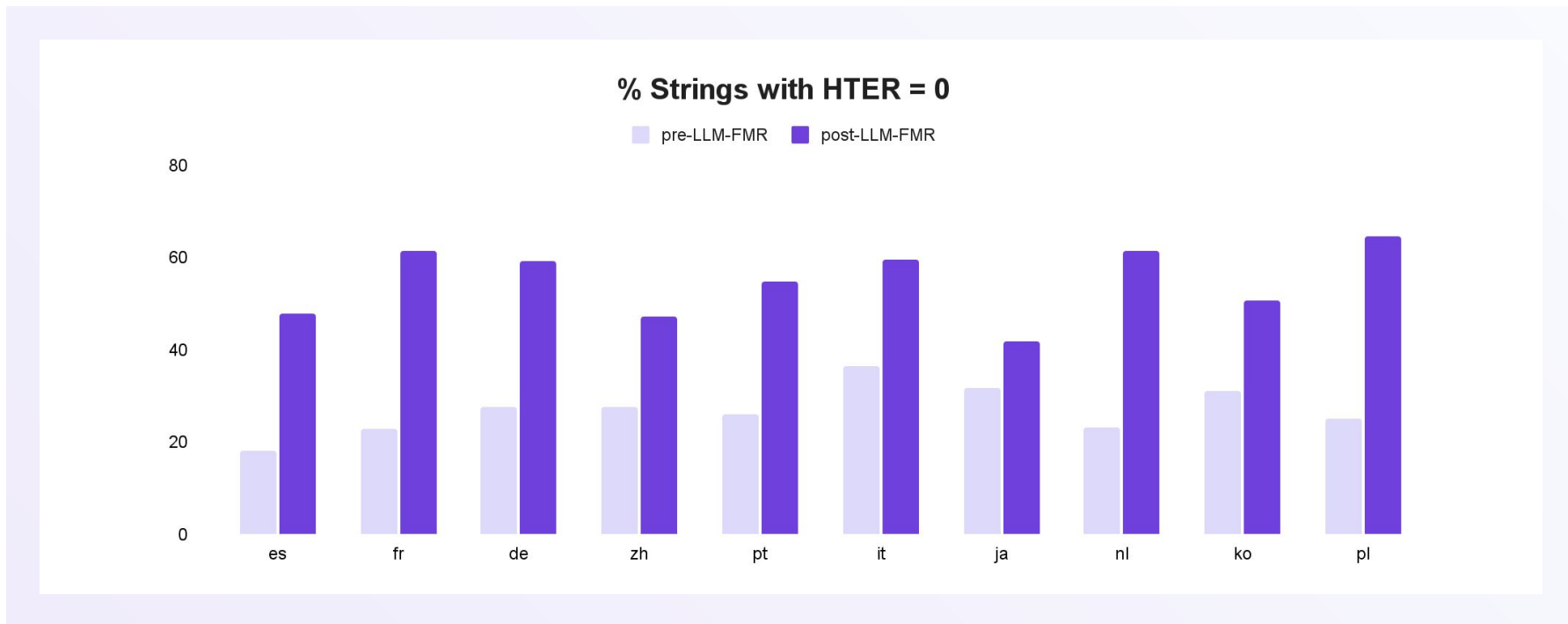
Without LLM	With LLM
Programe su **primer** **publicación**	Programe su **primera** **publicación**
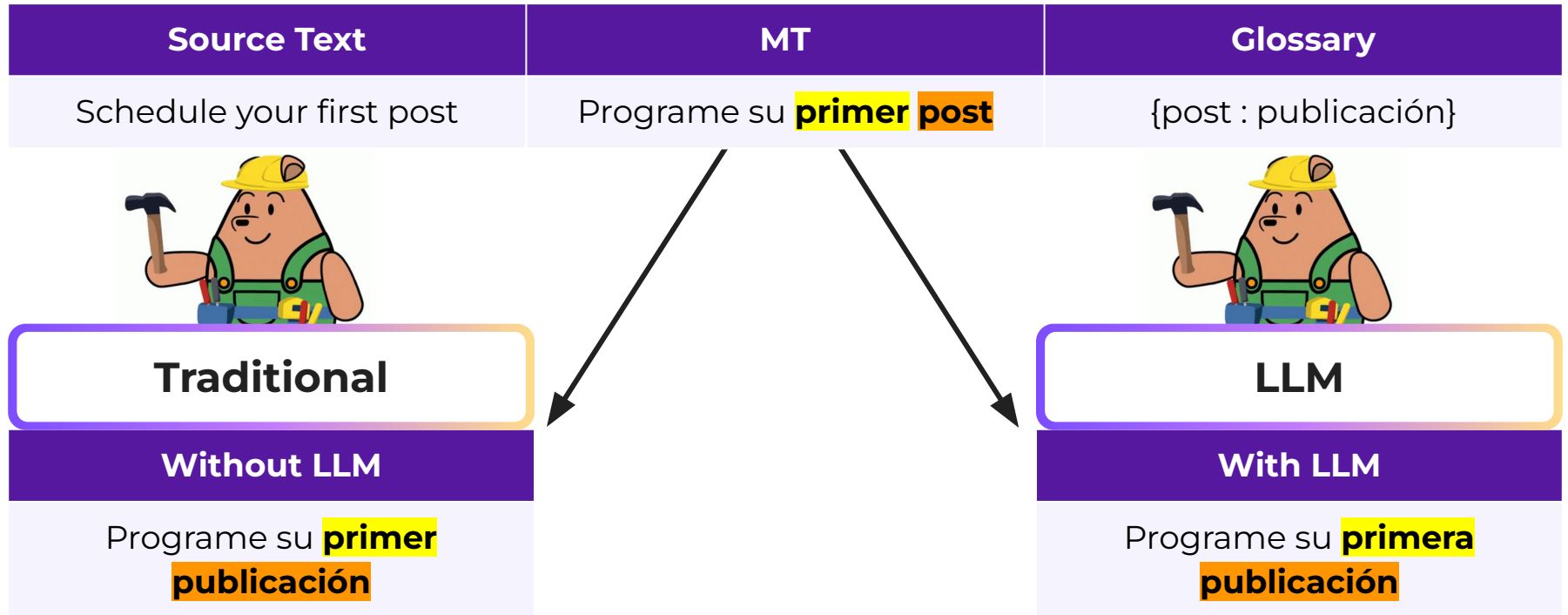
Glossary insertion: Case in point

Coverage

% Strings containing target terms **without** glossary insertion	% Strings containing target terms **with** glossary insertion
67%	96%

Quality

Locales	Without LLM	With LLM
	Morphologically correct insertions (%)	Morphologically correct insertions (%)
ru	62%	78%
es	80%	94%

*Quality is assessed with human evaluators

Edit Effort Estimation

Fine-tuned EEE

+ 13% accuracy over general EEE

General EEE in AI toolkit

+ 40% accuracy over random

Random guess for level of effort

SMARTLING

#globalready

GRC

Formality register is a continuum

Beach casual ← **Register Continuum** → **Black tie**

	Domain	
User generated content	**Domain**	Technical/legal Documents
Families, young kids, lower ranks	**Speaker and listener relation**	Strangers, older adults, higher ranks
Phone text messages	**Mode of communication**	Letters by mail
To connect to your end users	**Intention**	To promote high-end/luxury products/brands
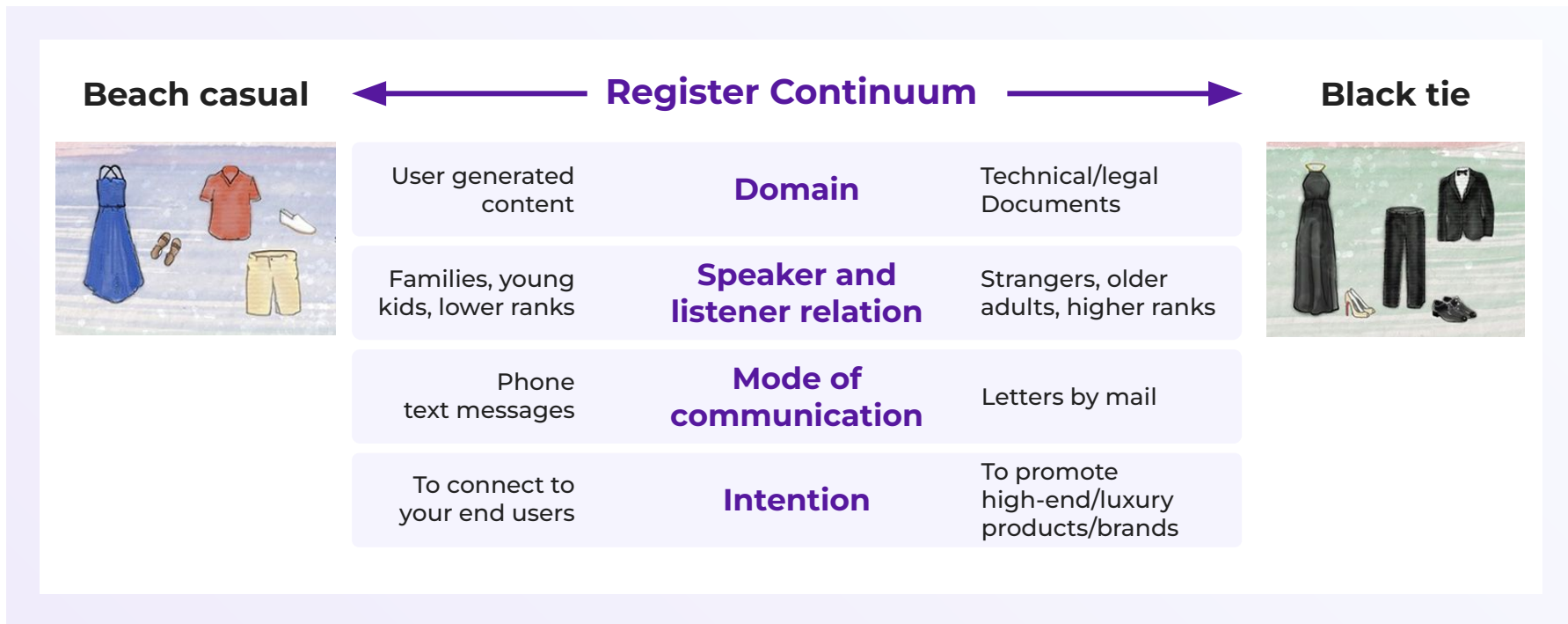
LLM for minimalistic register conversion

Task: Register Conversion (Formal to Informal) on PaLM: completion-bison

Success Criteria:
Convert formal strings to the informal register:

- With proper conjugations of the affected verbs, adjectives, nouns, etc.

- Without changing the tone and other word choices in the translation to make it overly informal.

Locales	% successful conversion	Remaining Issues
es-ES	95%	• Strings containing self-reflective verbs are not properly converted • Gender of original string not respected
de-DE	85%	• Imperatives with the second person pronoun in the beginning of the string are not properly converted while the rest of the strings are, resulting in mixed formality
fr-FR	80%	• Imperatives with the second person pronoun in the beginning of the string are not properly converted • Many strings became too informal due to truncation of the second person pronoun and the associated verb

SMARTLING   Research revealed: The impact of AI on translation   #globalready   GRC