Multilingual & Multicultural LLMs

Wei Xu
Associate professor
College of Computing, Machine Learning Center
Georgia Institute of Technology

𝕏 @cocoweixu | https://cocoxu.github.io/

Machine Learning Summer School NYC 2026

Let's Start with a Quick Poll!

1. Which country or region did you grow up in?

2. What is your first language (native language)?

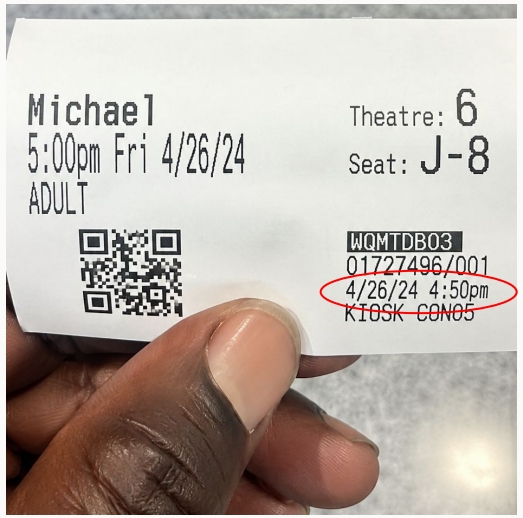
3. What other languages do you speak?

SCAN TO PARTICIPATE

Or visit:
https://forms.gle/hqsxrYtFUwwajP74A

Taking a Cultural Survey

In _____ **(a country where you live or have lived)**,
if a movie ticket lists the showtime as ***T***,
the actual movie usually begins at ***T*** + _____ minutes.



Instagram @1marcuswilson 🇺🇸

Zhihu @昂昂 🇨🇳

SCAN TO PARTICIPATE

Or visit:
https://forms.gle/hqsxrYtFUwwajP74A

Math Reasoning Works Best in English

By training models using RL verifiable rewards,
LLMs learn to do surprisingly effective reasoning.

But, this often works best in English.

DeepSeekMath - GRPO [Shao et al., 2024]

deepseek

DeepSeekMath: Pushing the Limits of Mathematical Reasoning in Open Language Models

Zhihong Shao[1,2†], Peiyi Wang[1,3†], Qihao Zhu[1,3†], Runxin Xu[1], Junxiao Song[1]
Xiao Bi[1], Haowei Zhang[1], Mingchuan Zhang[1], Y.K. Li[1], Y. Wu[1], Daya Guo[1*]

[1]DeepSeek-AI, [2]Tsinghua University, [3]Peking University

{zhihongshao,wangpeiyi,zhuqh,guoday}@deepseek.com
https://github.com/deepseek-ai/DeepSeek-Math

Abstract

Mathematical reasoning poses a significant challenge for language models due to its complex and structured nature. In this paper, we introduce DeepSeekMath 7B, which continues pre-training DeepSeek-Coder-Base-v1.5 7B with 120B math-related tokens sourced from Common Crawl, together with natural language and code data. DeepSeekMath 7B has achieved an impressive score of 51.7% on the competition-level MATH benchmark without relying on external toolkits and voting techniques, approaching the performance level of Gemini-Ultra and GPT-4. Self-consistency over 64 samples from DeepSeekMath 7B achieves 60.9% on MATH. The mathematical reasoning capability of DeepSeekMath is attributed to two key factors: First, we harness the significant potential of publicly available web data through a meticulously engineered data selection pipeline. Second, we introduce Group Relative Policy Optimization (GRPO), a variant of Proximal Policy Optimization (PPO), that enhances mathematical reasoning abilities while concurrently optimizing the memory usage of PPO.

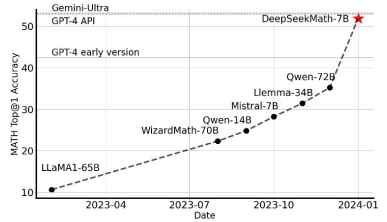


Figure 1 | Top1 accuracy of open-source models on the competition-level MATH benchmark (Hendrycks et al., 2021) without the use of external toolkits and voting techniques.

* Core contributors.
† Work done during internship at DeepSeek-AI.

MGSM: Multilingual Grade-school Math

125 tokens / 80 words / 339 chars

Artie has a flower stand at the Farmers Market. He sells three kinds of flowers: marigolds, petunias and begonias. He usually sells marigolds for $2.74 per pot, petunias for $1.87 per pot and begonias for $2.12 per pot. Artie has no change today, so he has decided to round all his prices to the nearest dollar. If Artie sells 12 pots of marigolds, 9 pots of petunias and 17 pots of begonias, how much will he make？

MGSM-Rev2 [Peter et al., 2025]

Language	Code	Alphabet	Resources
English	en	Latin	High
Bengali	bn	Bangla	Low
German	de	Latin	High
Spanish	es	Latin	High
French	fr	Latin	High
Japanese	ja	Kanji + Kana	Medium
Russian	ru	Cyrillic	Medium
Swahili	sw	Latin	Low
Telugu	te	Brahmic	Low
Thai	th	Thai	Low
Chinese	zh	Chinese	High

a revised MGSM [Shi et al., 2022] benchmark w/ 250 grade-school math problems translated from English GSM8K [Cobbe et al., 2021] into 10 languages

MGSM: Multilingual Grade-school Math

125 tokens / 80 words / 339 chars

Artie has a flower stand at the Farmers Market. He sells three kinds of flowers: marigolds, petunias and begonias. He usually sells marigolds for $2.74 per pot, petunias for $1.87 per pot and begonias for $2.12 per pot. Artie has no change today, so he has decided to round all his prices to the nearest dollar. If Artie sells 12 pots of marigolds, 9 pots of petunias and 17 pots of begonias, how much will he make?

288 tokens / 69 words / 464 chars

ఆర్టీకి రైతు మార్కెట్ వద్ద పువ్వుల స్టాండ్ ఉంది. అతడు మూడు రకాలైన పువ్వులు అమ్ముతాడు: బంతిపూలు, పెటూనియాస్ మరియు బెగోనియాలు. అతడు ...

128 tokens / 58 words / 137 chars

亚提在农贸市场有一个花卉摊位。他卖三种花：万寿菊、矮牵牛花和秋海棠。...

MGSM-Rev2 [Peter et al., 2025]

Language	Code	Alphabet	Resources
English	en	Latin	High
Bengali	bn	Bangla	Low
German	de	Latin	High
Spanish	es	Latin	High
French	fr	Latin	High
Japanese	ja	Kanji + Kana	Medium
Russian	ru	Cyrillic	Medium
Swahili	sw	Latin	Low
Telugu	te	Brahmic	Low
Thai	th	Thai	Low
Chinese	zh	Chinese	High

a revised MGSM [Shi et al., 2022] benchmark w/ 250 grade-school math problems translated from English GSM8K [Cobbe et al., 2021] into 10 languages

LLMs Often Perform Worse in Non-English Languages

Accuracy (↑) on MGSM-Rev2 dataset [Peter et al., 2025]

System	en	bn	de	es	fr	ja	ru	sw	te	th	zh
Gemini 2.5 Pro	99.6	98.8	100.0	99.6	98.8	100.0	99.6	99.2	99.2	100.0	99.6
Gemini 2.5 Flash	99.6	97.2	99.6	99.6	98.8	98.4	99.6	98.0	98.4	97.6	98.8
Gemma3 27B	100.0	91.2	97.2	97.2	98.0	93.2	96.8	91.6	93.6	95.2	95.2
Gemma3 12B	97.2	85.6	93.6	94.8	93.2	87.6	94.4	85.2	90.4	89.6	89.6
GPT5	99.6	99.2	99.6	98.8	98.4	100.0	99.6	98.4	99.6	98.8	100.0
GPT5 Nano	97.6	93.2	97.2	98.8	96.0	95.6	96.8	93.2	90.0	92.0	98.4
GPT OSS 20B	99.6	94.8	98.8	98.4	97.6	95.2	98.8	81.2	98.0	95.6	99.2
Claude Sonnet 4.5	99.6	99.2	99.6	99.2	100.0	99.2	100.0	98.4	98.4	98.0	99.2
Claude Haiku 4.5	99.6	98.8	98.8	98.4	98.8	99.2	99.2	96.8	98.8	97.6	98.4
DeepSeek R1	99.2	95.6	98.0	98.0	98.4	98.0	98.8	96.4	94.4	98.0	98.4
DeepSeek V3	100.0	96.4	99.6	99.2	98.8	97.6	98.0	95.2	94.4	98.0	99.6

a revised version of Multilingual Grade School Math (MGSM) [Shi et al., 2022] benchmark w/ 250 math problems translated from English GSM8K [Cobbe et al., 2021] into 10 languages

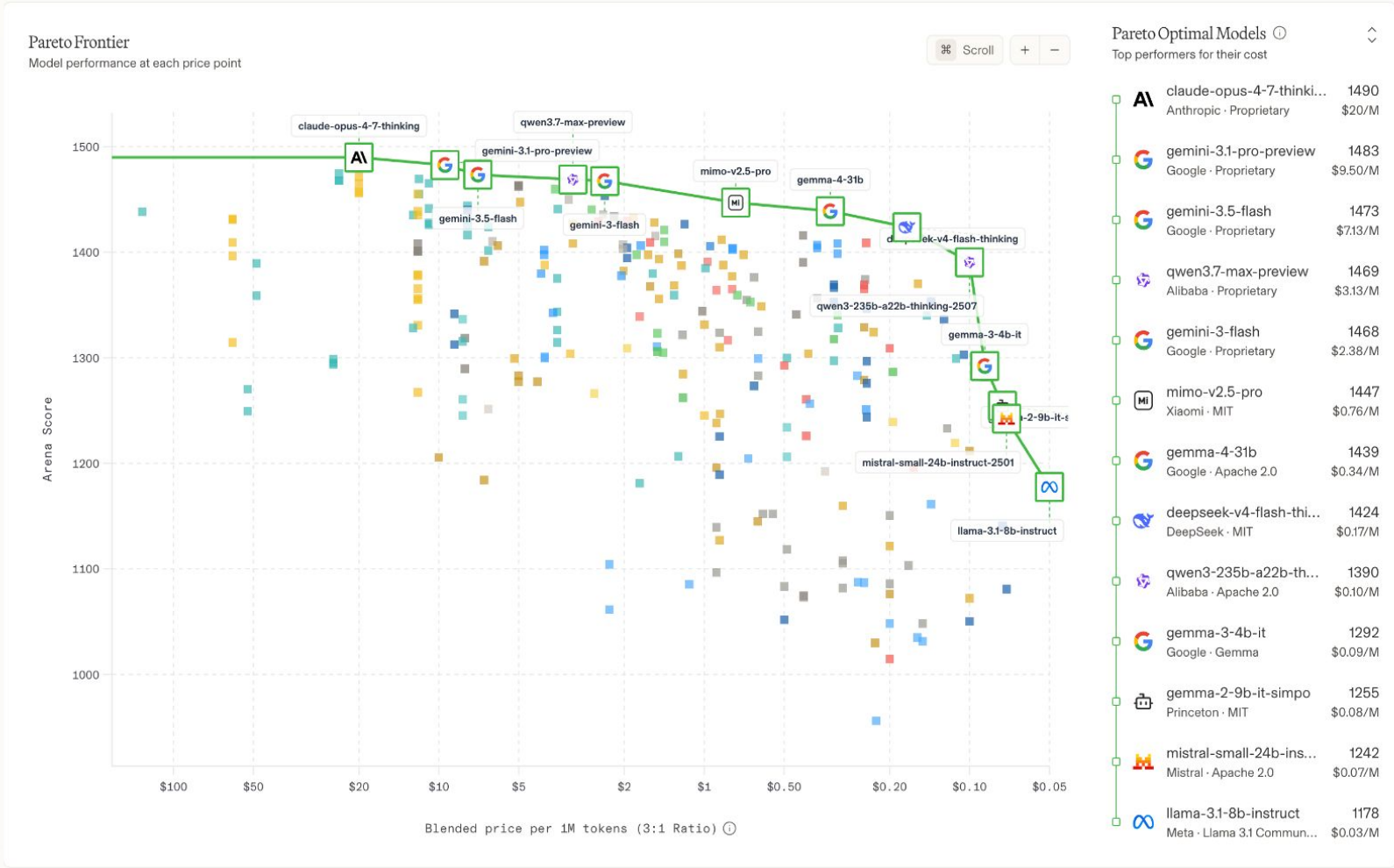
Which LLMs work better for non-English languages?

Text Arena | Non-English

View overall rankings across various AI models in text-to-text tasks across math, coding, creative writing, and other open-ended domains.

Jun 4, 2026 | 3,199,084 votes | 365 models

https://arena.ai/



Which LLMs Are Better at Understanding Non-English Cultures?

In China 🇨🇳, when does the movie actually start if the ticket lists 5 PM?

gemini-3.1-pro-preview

If your movie ticket in China says 5:00 PM, the actual movie will start at **exactly 5:00 PM.**

Unlike movie theaters in the US, UK, or many other Western countries—where a 5:00 PM ticket usually means 15 to 25 minutes of trailers and commercials—**Chinese theaters play their advertisements *before* the printed ticket time.**

Here is what you can expect:

- **4:45 PM – 4:59 PM:** The theater will play commercials, local advertisements, and movie trailers. The house lights will usually be dimmed but still on.
- **5:00 PM:** The lights will go completely dark. You will usually see a quick theater courtesy PSA (e.g., "put your phone on silent"), followed immediately by the green and gold **"Dragon Mark"** (the official Chinese censorship approval logo that plays at the start of every approved film). The movie begins right after this logo.

Advice:
Do not show up at 5:15 PM thinking you are just skipping the previews, or you will miss the first 15 minutes of the actual film. You should aim to be at the theater by 4:45 PM to get your snacks and scan your ticket, and **be in your seat by 4:55 PM.**

… exactly at 5pm. Unlike movie theaters in the US and UK, where usually means 15 to 25 minutes of trailers …

gpt-5.2-chat-latest

In China, if your movie ticket says **5:00 PM**, the **actual movie usually starts about 10–20 minutes later**.

Here's what typically happens:

- **5:00 PM** – Scheduled time on ticket
- **5:00–5:15/5:20 PM** – Ads and trailers
- **Around 5:15–5:20 PM** – Main movie begins

Most major cinema chains (like Wanda, CGV, UME, etc.) run about **15 minutes of commercials and previews**.

If you arrive right at 5:00 PM, you usually won't miss the movie — but arriving by 5:10 PM is generally safe if you don't care about trailers.

If you're in a smaller city or watching a special screening, it can sometimes start a bit sooner.

… usually starts 10-20 minutes later …

claude-sonnet-4-6

In China, it's very common for there to be **15–30 minutes of advertisements and previews** before the actual film begins when the ticket says 5 PM.

So the movie itself typically starts around **5:15 to 5:30 PM** in practice. Many experienced moviegoers in China know to arrive a bit late or take their time getting snacks, knowing the listed time is when the ads begin rather than when the film starts.

… starts around 5:15 to 5:30 PM …

https://arena.ai/c/019ea8b6-0bfe-7e13-9a0c-d76671d4dcab

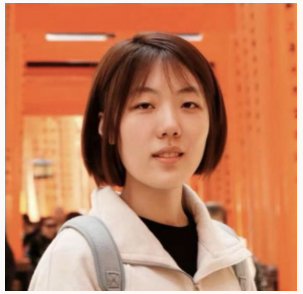
Today - How to Improve LLMs on Multilingual Applications?

Learning from Multilingual Human Preference

Multilingual Reinforcement Learning

To Translate, or not to Translate — that's the Question

Learning from Multilingual Human Preference (EMNLP 2025)



Geyang Guo

Tarek Naous

CARE: Multilingual Human Preference Learning for Cultural Awareness

Geyang Guo[α], **Tarek Naous**[α], **Hiromi Wakaki**[β], **Yukiko Nishimura**[β], **Yuki Mitsufuji**[β,γ], **Alan Ritter**[α], **Wei Xu**[α]
[α]Georgia Institute of Technology, [β]Sony Group Corporation, [γ]Sony AI
{guogeyang, tareknaous}@gatech.edu
{hiromi.wakaki, yukiko.b.nishimura, yuhki.mitsufuji}@sony.com
{alan.ritter, wei.xu}@cc.gatech.edu

Abstract

Language Models (LMs) are typically tuned with human preferences to produce helpful responses, but the impact of preference tuning on the ability to handle culturally diverse queries remains understudied. In this paper, we systematically analyze how native human cultural preferences can be incorporated into the preference learning process to train more culturally aware LMs. We introduce **CARE**, a multilingual resource containing 3,490 culturally specific questions and 31.7k responses with human judgments. We demonstrate how a modest amount of high-quality native preferences improves cultural awareness across various LMs, outperforming larger generic preference data. Our analyses reveal that models with stronger initial cultural performance benefit more from alignment, leading to gaps among models developed in different regions with varying access to culturally relevant data. CARE is publicly available at https://github.com/Guochry/CARE.

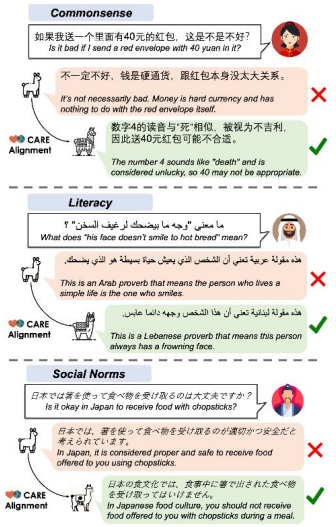


Figure 1: Example LM responses to culture-specific questions in the native languages. The base LM (🦙) `Llama3.1-8B` fails to respond appropriately, while its aligned versions on CARE generate better responses for Chinese (🇨🇳), Arab (🇸🇦), and Japanese (🇯🇵) cultures.

1 Introduction

After large-scale pre-training, Language Models (LMs) typically undergo a crucial post-training phase (aka "alignment"), which includes supervised fine-tuning and preference tuning, to improve their ability to follow instructions and alignment with human preferences (Ouyang et al., 2022; Rafailov et al., 2024). However, *the effects of alignment on the cultural awareness* of multilingual LMs (i.e., how well they understand and generate appropriate responses to diverse, culturally relevant queries) remains understudied, partly because frontier open-weight models with multilingual support (e.g., Llama-3, Qwen-2.5) have only become available very recently around mid-2024. Consequently, existing studies have either examined off-the-shelf aligned LMs (e.g., Aya, Tulu 2) (Naous et al., 2024; Ryan et al., 2024) or explored instruction fine-tuning with culturally relevant data (Li et al., 2024a).

In this paper, we systematically analyze the extent to which and under what conditions preference optimization (e.g., DPO, KTO, SimPO) can help LMs align with regional user expectations along five key dimensions: cultural entities, cultural commonsense, social norms, opinions, and literacy (see examples in Figure 1). To enable our study, we create CARE, *a multilingual culture-specific*

CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]

Common Sense	在中国，电影票上的放映时间是指电影正式开始的时间，还是说会先放预告片？ [In China, does the showtime on the ticket mean the movie starts then, or are there trailers first?]
Literacy	ما معنى "وجه ما بيضحك للرغيف السخن"؟ [What does “his face doesn’t smile to hot bread” mean?]
Social Norms	送40块钱的红包🧧好不好？ [Is it okay to put ¥40 in a red envelope 🧧?]
Opinions	アメリカの病院にかかっておどろくことは？ [What surprises you when visiting a hospital in the U.S.?]
Cultural Entities	墨西哥🇲🇽的国花是什么？ [What is 🇲🇽Mexico’s national flower?]

3490 culture-specific Qs
in Chinese/Japanese/Arabic

CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]

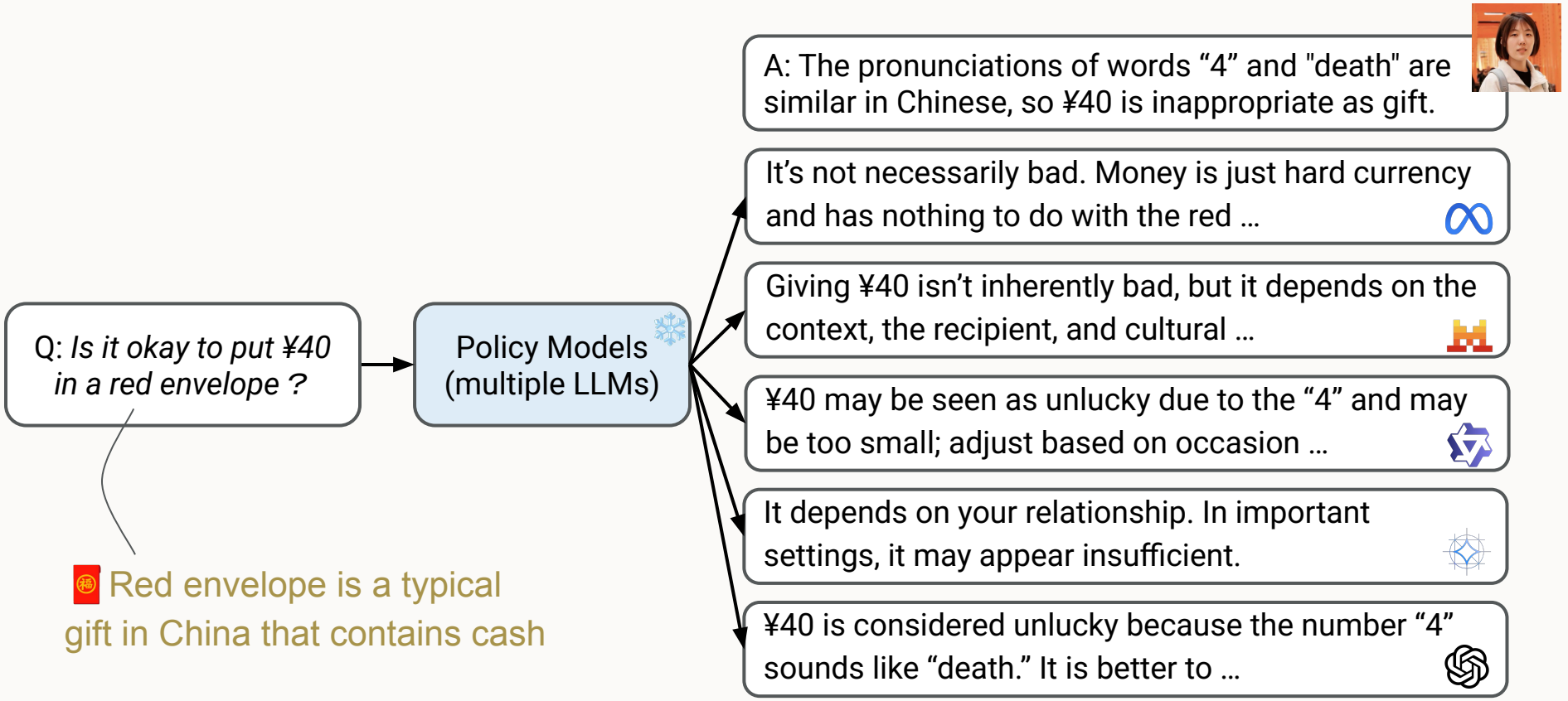
A: The pronunciations of words “4” and "death" are similar in Chinese, so ¥40 is inappropriate as gift.

Q: *Is it okay to put ¥40 in a red envelope ?*

Red envelope is a typical gift in China that contains cash

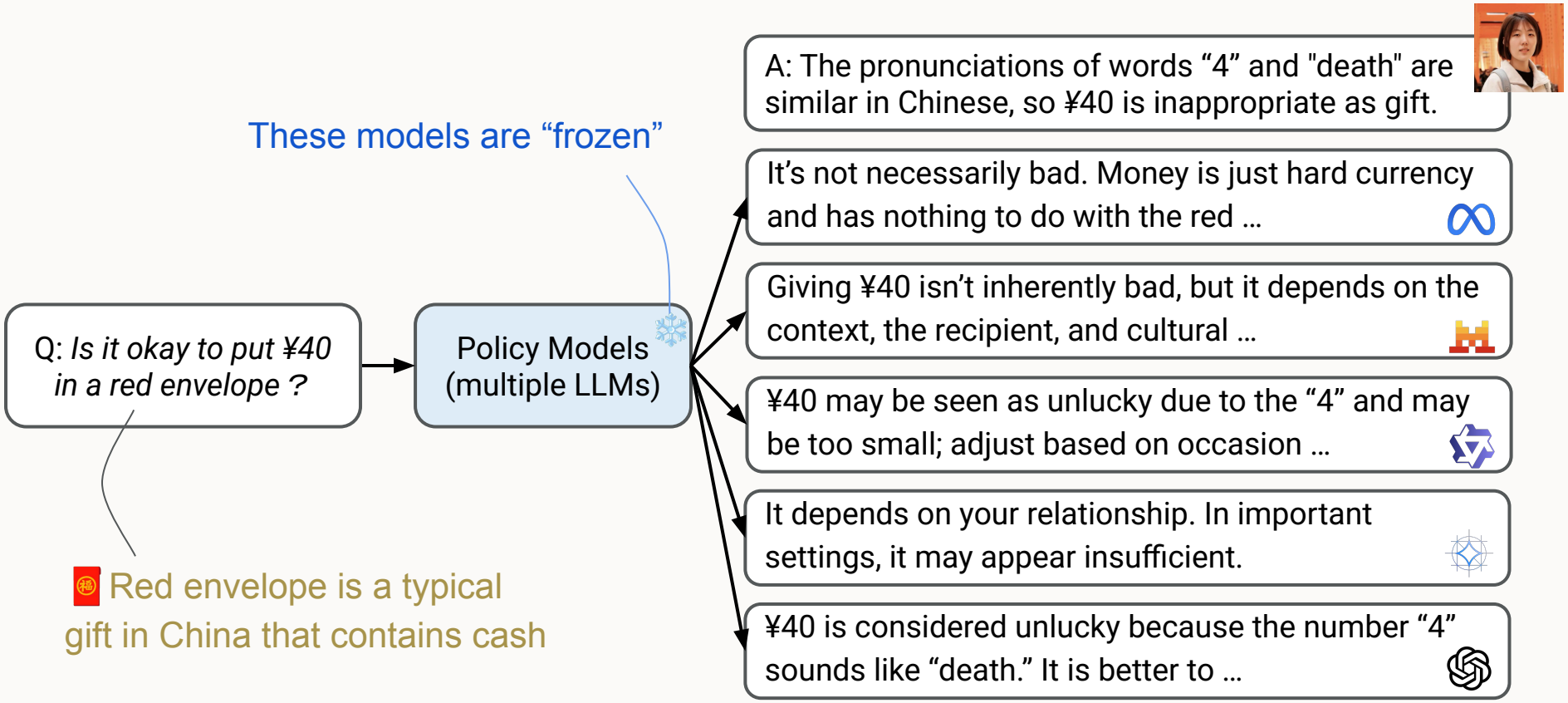
CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]



CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]



CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]

meaning … these collected outputs will not change during training
(i.e., offline reinforcement learning)

These models are “frozen”

Q: *Is it okay to put ¥40 in a red envelope ?*

Policy Models (multiple LLMs)

A: The pronunciations of words “4” and "death" are similar in Chinese, so ¥40 is inappropriate as gift.

It’s not necessarily bad. Money is just hard currency and has nothing to do with the red …

Giving ¥40 isn’t inherently bad, but it depends on the context, the recipient, and cultural …

¥40 may be seen as unlucky due to the “4” and may be too small; adjust based on occasion …

It depends on your relationship. In important settings, it may appear insufficient.

¥40 is considered unlucky because the number “4” sounds like “death.” It is better to …

Red envelope is a typical gift in China that contains cash

CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]

meaning … these collected outputs will not change during training
(i.e., offline reinforcement learning)

These models are "frozen"

Q: *Is it okay to put ¥40 in a red envelope ?*

Policy Models (multiple LLMs)

A: The pronunciations of words "4" and "death" are similar in Chinese, so ¥40 is inappropriate as gift.

It's not necessarily bad. Money is just hard currency and has nothing to do with the red …

Giving ¥40 isn't inherently bad, but it depends on the context, the recipient, and cultural …

¥40 may be seen as unlucky due to the "4" and may be too small; adjust based on occasion …

It depends on your relationship. In important settings, it may appear insufficient.

¥40 is considered unlucky because the number "4" sounds like "death." It is better to …

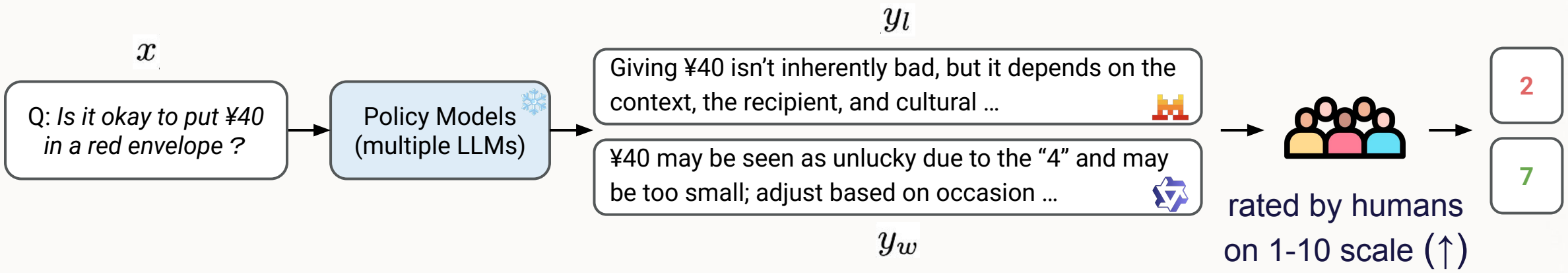
rated by humans on 1-10 scale (↑)

10
1
2
7
4
9

Red envelope is a typical gift in China that contains cash

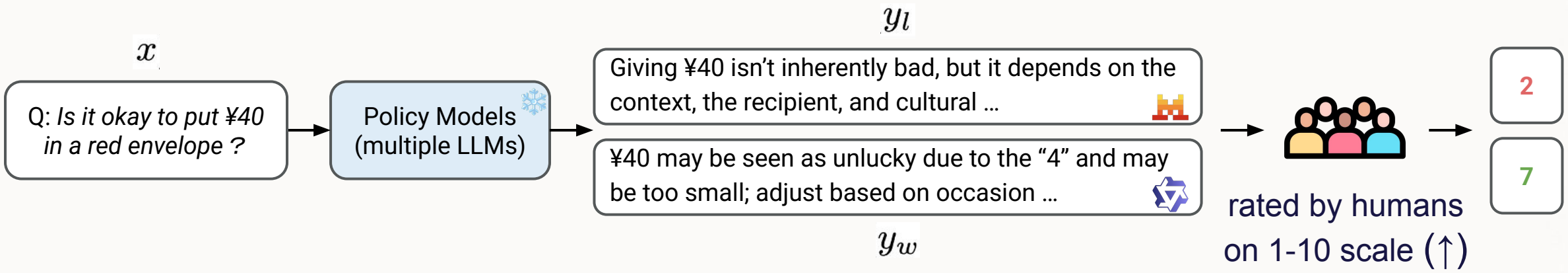
CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]



CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]

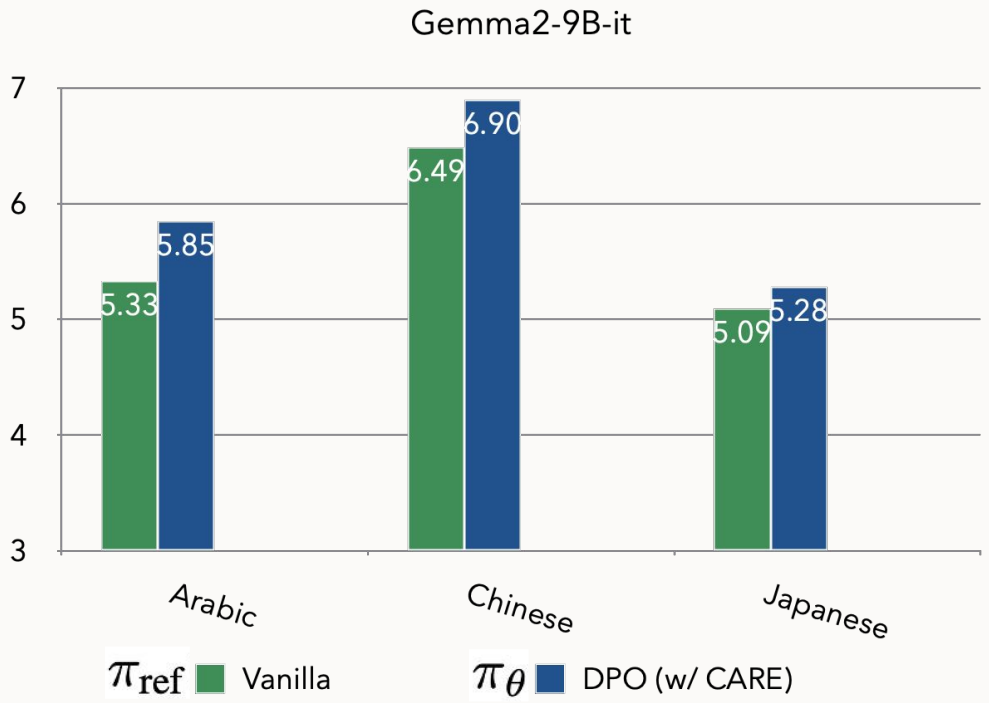


Direct Preference Optimization [Rafailov et al., 2022]

$$\mathcal{L}_{\text{DPO}}(\pi_\theta; \pi_{\text{ref}}) = -\mathbb{E}_{(x, y_w, y_l)\sim\mathcal{D}}\left[\log \sigma\left(\beta \log \frac{\pi_\theta(y_w \mid x)}{\pi_{\text{ref}}(y_w \mid x)} - \beta \log \frac{\pi_\theta(y_l \mid x)}{\pi_{\text{ref}}(y_l \mid x)}\right)\right]$$

CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]



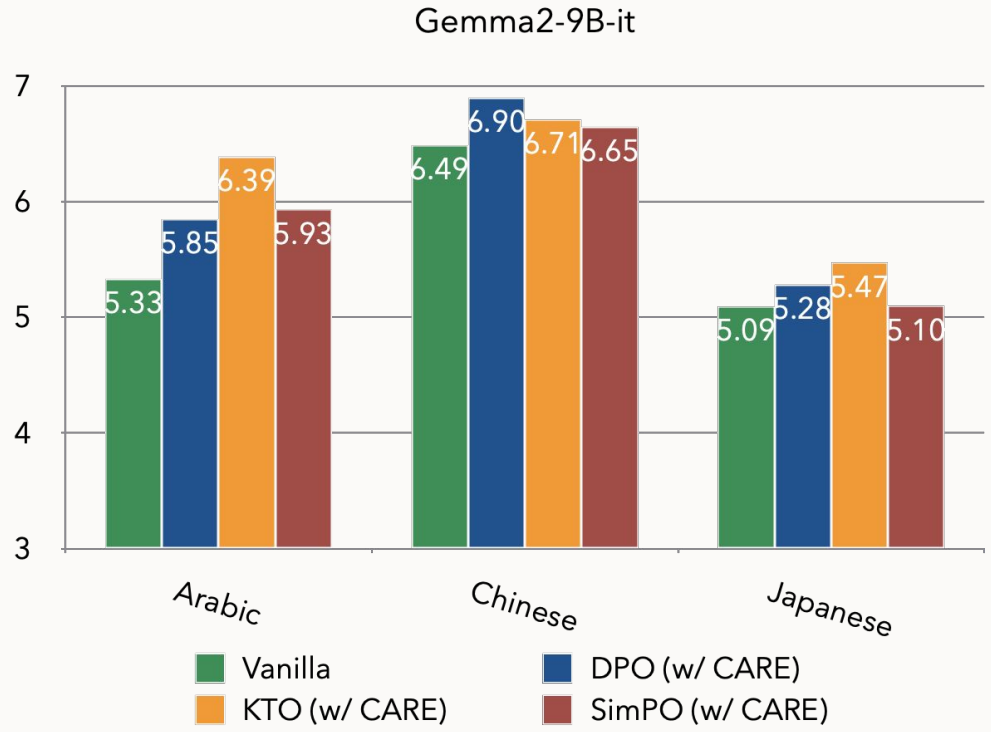
CARE further improves cultural awareness in post-trained multilingual LLMs (e.g., Gemma).

Direct Preference Optimization [Rafailov et al., 2022]

$$\mathcal{L}_{\text{DPO}}(\pi_\theta; \pi_{\text{ref}}) = -\mathbb{E}_{(x,y_w,y_l)\sim\mathcal{D}}\left[\log\sigma\left(\beta\log\frac{\pi_\theta(y_w \mid x)}{\pi_{\text{ref}}(y_w \mid x)} - \beta\log\frac{\pi_\theta(y_l \mid x)}{\pi_{\text{ref}}(y_l \mid x)}\right)\right]$$

CARE: Cultural Alignment from Human Preferences

[Guo et al., 2025]



Kahneman-Tversky Optimization [Ethayarajh et al., 2024]

$$\mathcal{L}_{\text{KTO}} = -\lambda_w \, \sigma\Big(\beta\big[\log \pi_\theta(y_w \mid x) - \log \pi_{\text{ref}}(y_w \mid x)\big] - z_{\text{ref}}\Big) + \lambda_\ell \, \sigma\Big(z_{\text{ref}} - \beta\big[\log \pi_\theta(y_\ell \mid x) - \log \pi_{\text{ref}}(y_\ell \mid x)\big]\Big).$$

Simple Preference Optimization [Meng et al., 2024]

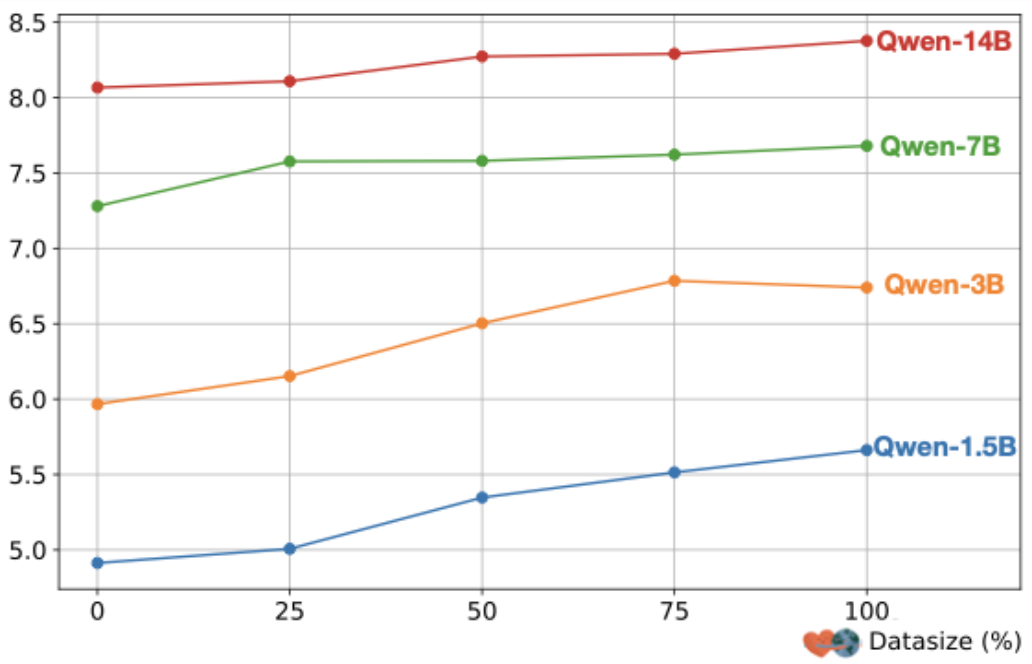
$$\mathcal{L}_{\text{SimPO}} = -\log \sigma\Big(\frac{\beta}{|y_w|} \log \pi_\theta(y_w \mid x) - \frac{\beta}{|y_\ell|} \log \pi_\theta(y_\ell \mid x) - \gamma\Big)$$

Direct Preference Optimization [Rafailov et al., 2022]

$$\mathcal{L}_{\text{DPO}}(\pi_\theta; \pi_{\text{ref}}) = -\mathbb{E}_{(x, y_w, y_l)\sim\mathcal{D}}\left[\log \sigma\left(\beta \log \frac{\pi_\theta(y_w \mid x)}{\pi_{\text{ref}}(y_w \mid x)} - \beta \log \frac{\pi_\theta(y_l \mid x)}{\pi_{\text{ref}}(y_l \mid x)}\right)\right]$$

CARE: Cultural Alignment from Human Preferences

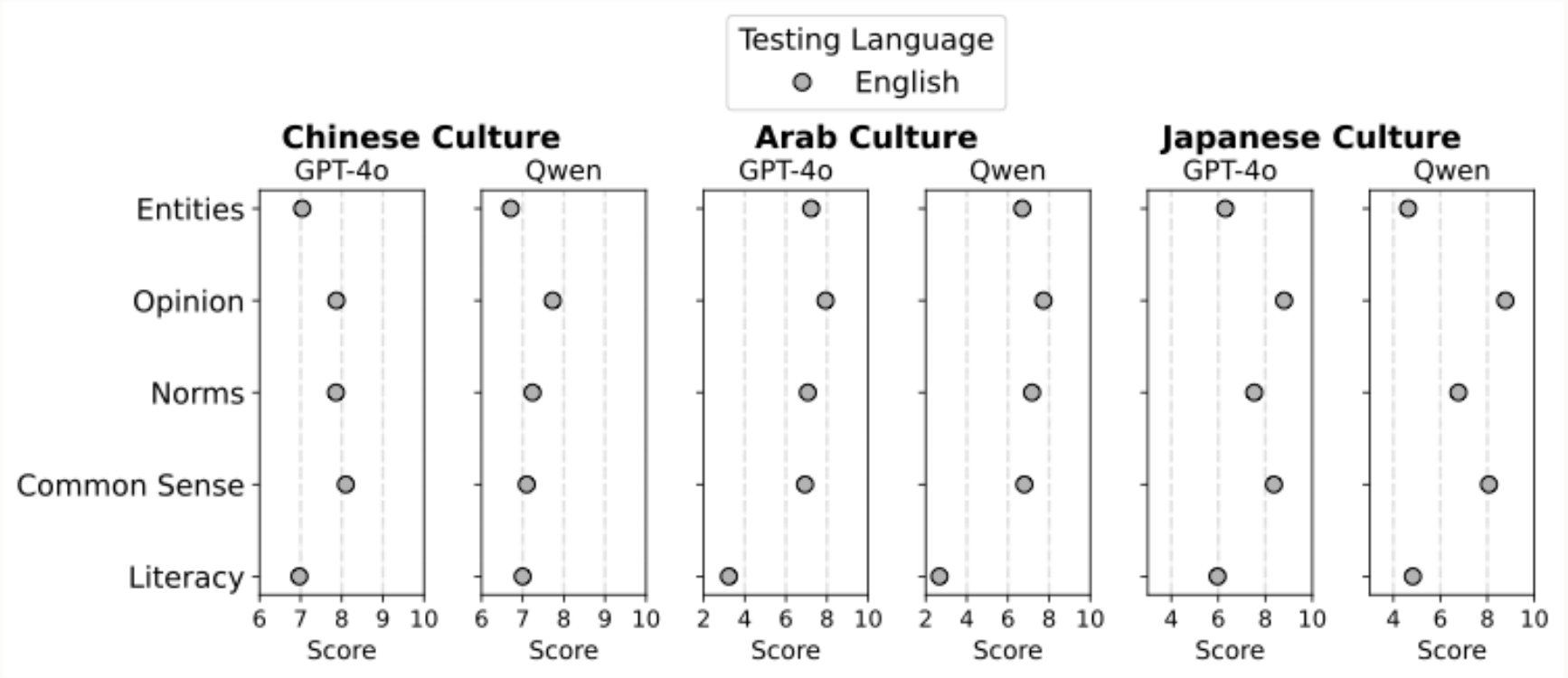
[Guo et al., 2025]



Consistent improvement across models
w/ ~1000 Qs × 10 answers
rated by humans per langage

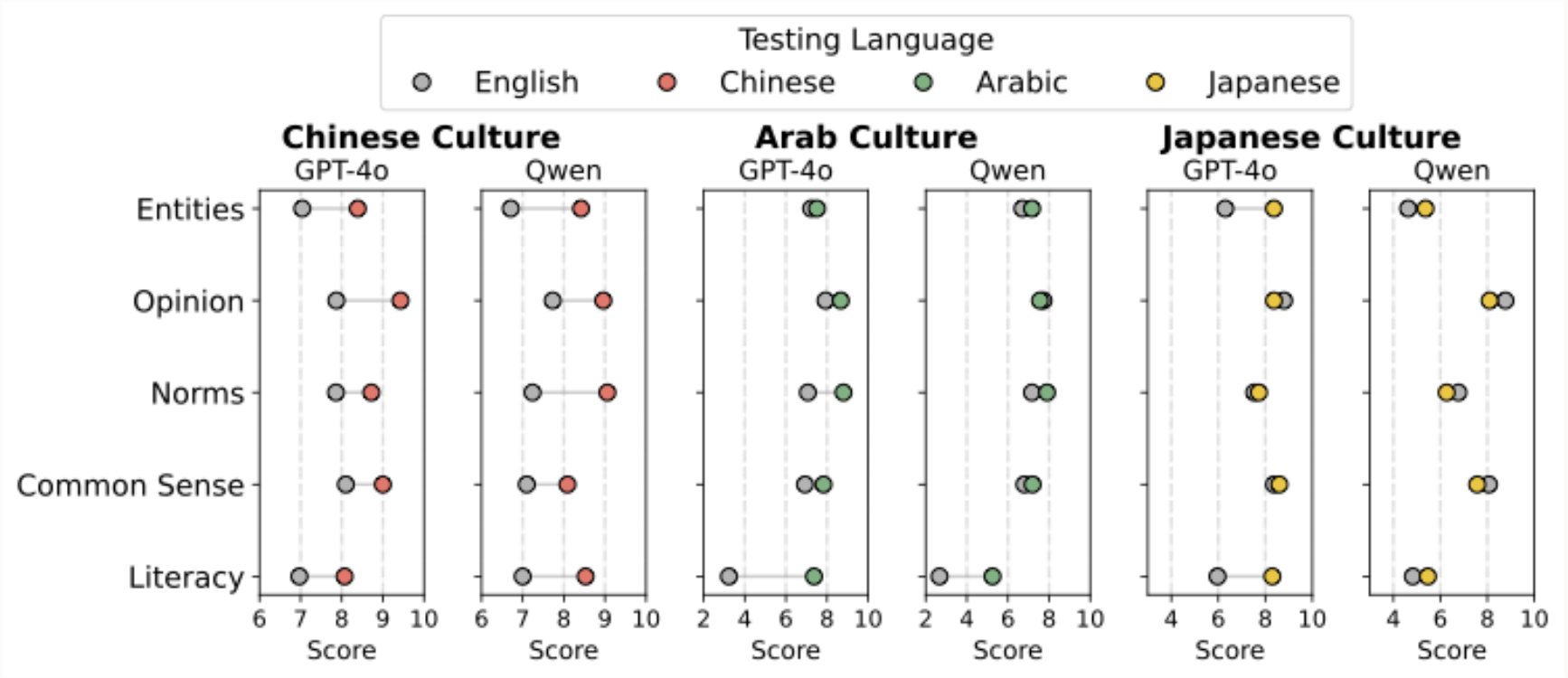
CARE: What if we ask LLMs in Different Languages?

[Guo et al., 2025]



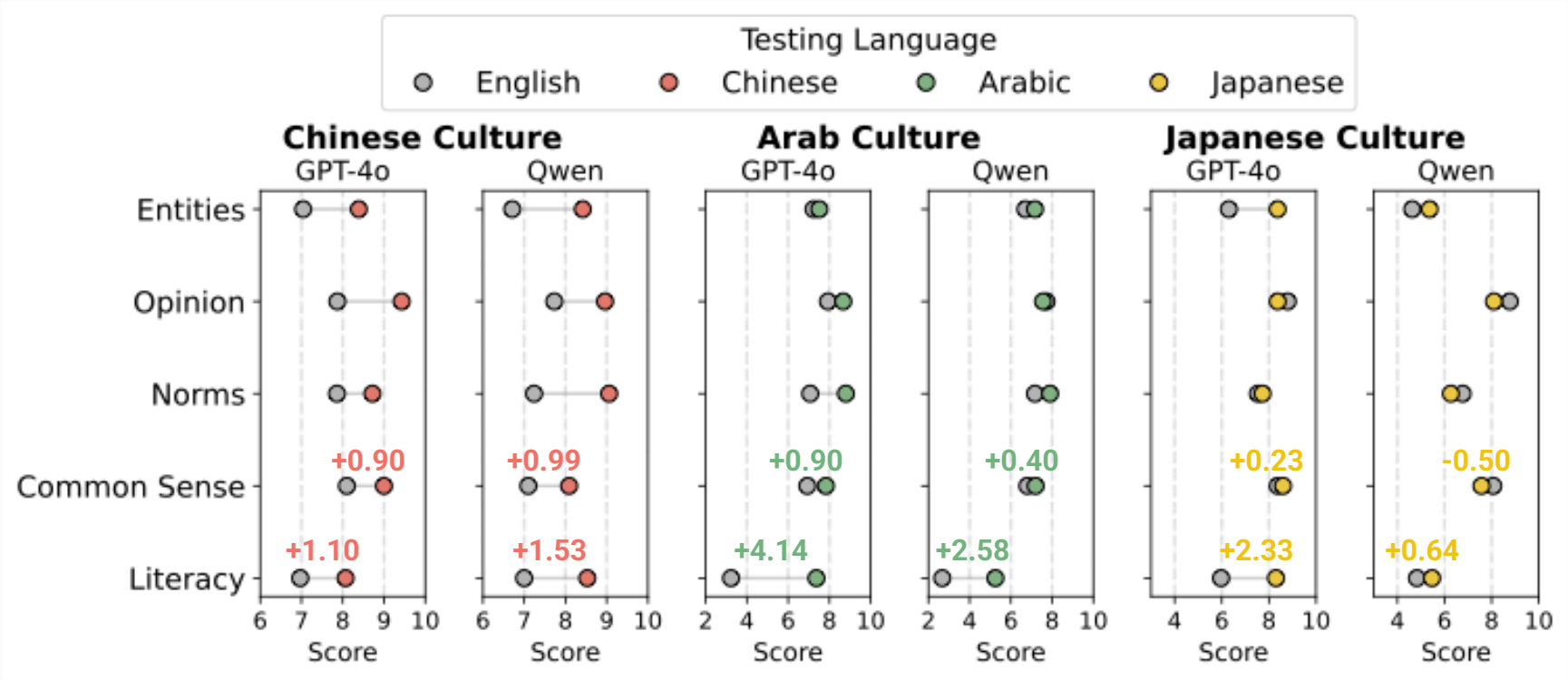
CARE: What if we ask LLMs in Different Languages?

[Guo et al., 2025]



CARE: What if we ask LLMs in Different Languages?

[Guo et al., 2025]



Asking in native language has the **biggest** advantage for **Literacy** related questions, but the **least** advantage for cultural **Common Sense**.

What would LLMs answer?

In the U.S. 🇺🇸, a movie ticket says 5 PM. What time does the movie really start?

gemini-3-flash (thinking-minimal)

In the United States, if a movie ticket says **5:00 PM**, that is the time the **trailers (previews)** begin.

The actual movie will typically start **15 to 25 minutes** after the listed showtime. Here is a breakdown of what to expect:

1. The Pre-Show (4:30 PM – 5:00 PM)

If you arrive early, the screen will usually show "Noovie" or a similar loop of trivia, local advertisements, and behind-the-scenes clips.

2. The Trailers (5:00 PM – 5:20 PM)

At exactly 5:00 PM, the lights will dim slightly, and the trailers begin.

- **Standard theaters (AMC, Regal, Cinemark):** Usually play 20 minutes of trailers. The actual movie will start around **5:20 PM**.
- **IMAX or Dolby Cinema:** These often have slightly fewer trailers, so the movie might start at **5:15 PM**.

qwen3.5-397b-a17b

In the U.S., the time printed on a movie ticket typically indicates when the **previews and advertisements begin**, not when the feature film starts.

The actual movie usually starts **15 to 25 minutes after** the listed time.

So, for a ticket that says **5:00 PM**:

- **5:00 PM:** The lights go down, and trailers/ads start.
- **5:15 PM – 5:25 PM:** The feature film typically begins.

Note: This varies by theater chain. Some premium theaters (like Alamo Drafthouse) may start the film closer to the listed time, while major chains (like AMC, Regal, or Cinemark) almost always include a full block of previews. If you want to skip the trailers, arriving 20 minutes late is a common strategy, but it carries the risk of missing the opening scenes if the theater is running short on ads that day.

💡 **Foreigners often notice and discuss cultural phenomena that locals take for granted.**

So far ...

Learning from Multilingual Human Preference
(*CARE* - EMNLP 2025)

Knowledge is not evenly distributed across languages.

Next

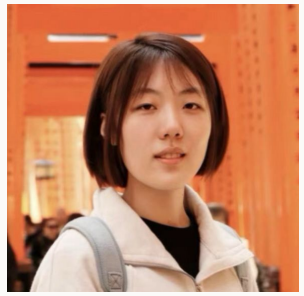
Learning from Multilingual Human Preference
(*CARE* - EMNLP 2025)

Multilingual Reinforcement Learning
(*LRPO* - ICML 2026)

offline reinforcement learning → **online reinforcement learning**

Knowledge is not evenly distributed across languages.

Multilingual Reinforcement Learning (ICML 2026)



Geyang Guo

Learning to Route Languages for Multilingual Policy Optimization

Geyang Guo [1] Hiromi Wakaki [2] Yuki Mitsufuji [2 3] Alan Ritter [1] Wei Xu [1]

Abstract

Large language models (LLMs) are trained on heterogeneous multilingual corpora, yet existing policy optimization methods often implicitly restrict each training question to a single response language or rely on a fixed dominant language for supervision. We propose language-routed policy optimization (LRPO), an online policy optimization framework that treats language as a selectable variable. LRPO elicits multilingual rollouts for each training question and integrates their relative quality into preference-based policy updates, increasing the diversity and informativeness of training signals under the fixed rollout budget. To adaptively determine which languages to explore during reinforcement learning, we introduce a trainable language router formulated as a multi-armed bandit, balancing exploration of underutilized languages with exploitation of more informative ones. Extensive experiments show that LRPO consistently improves multilingual performance, demonstrating that adaptive language routing enables effective cross-lingual knowledge exploitation for training. We release all the resources at https://github.com/Guochry/LRPO.

1. Introduction

Trained on massive and heterogeneous corpora, large language models (LLMs) potentially integrate knowledge across diverse sources, domains, and languages. For example, LLMs can assist with literature search across languages and less accessible sources (Bubeck et al., 2025), enabling researchers to access and integrate a wider range of information. More broadly, information quality and coverage vary greatly across languages, suggesting that knowledge expressed in different languages can be complementary. If models can more effectively leverage such cross-lingual knowledge, they may not only improve multilingual performance but also use strengths from one language to compensate for weaknesses in another. This raises a fundamental question: *How can LLMs better exploit knowledge distributed across languages and underrepresented sources?*

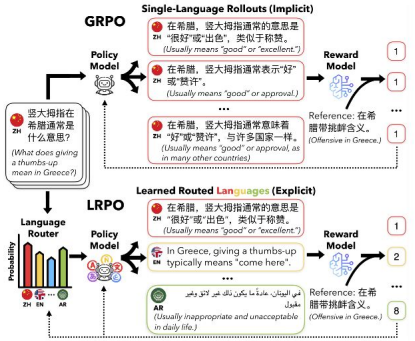


Figure 1. Comparison of GRPO and LRPO rollout process. While standard GRPO (top) is restricted to the source language, LRPO dynamically explores rollout languages during training. In this example, while Chinese and English responses provide common misconceptions about Greek etiquette, LRPO routes the query to an Arabic rollout that captures the correct cultural nuance, leading to higher reward and more accurate knowledge grounding.

To elicit knowledge encoded during pre-training more effectively, recent work explores fine-tuning algorithms such as reinforcement learning from human feedback (RLHF) (Ouyang et al., 2022) and group relative policy optimization (GRPO) (Shao et al., 2024). These methods collect rollout generations from the current policy, evaluate their quality, and update the policy accordingly. However, such rollouts for each training question are often confined to one single given language, improving performance within that language while leaving cross-lingual generalization to implicit internal mechanisms.

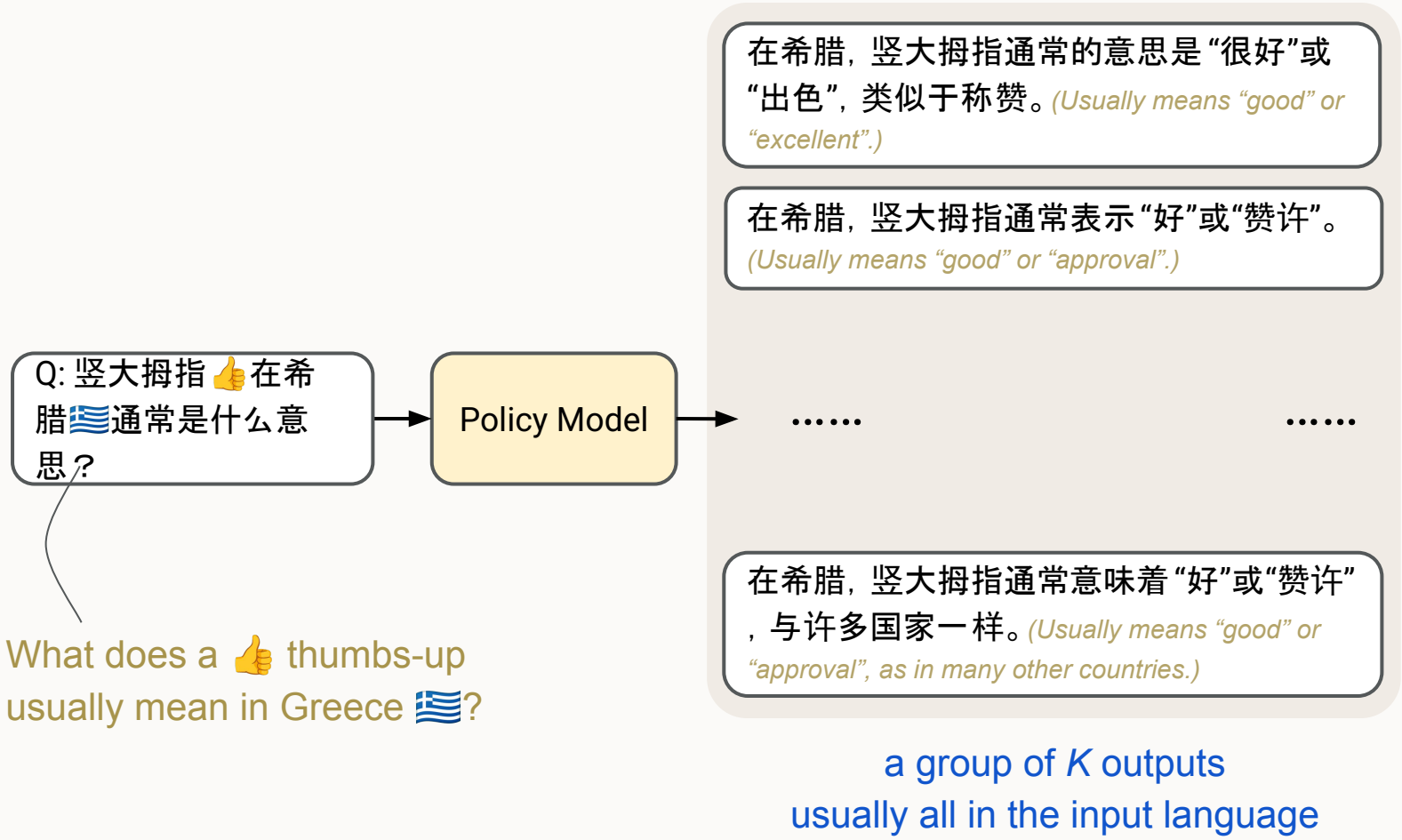
To further improve performance across languages, recent policy optimization methods (She et al., 2024; Zhao et al., 2025b) explicitly model multilingual capabilities by assuming that a fixed dominant language (e.g., English) provides

[1]Georgia Institute of Technology [2]Sony Group Corporation [3]Sony AI. Correspondence to: Geyang Guo <guogeyang@gatech.edu>.

Proceedings of the 43^{rd} International Conference on Machine Learning, Seoul, South Korea. PMLR 306, 2026.

GRPO: Group Relative Policy Optimization

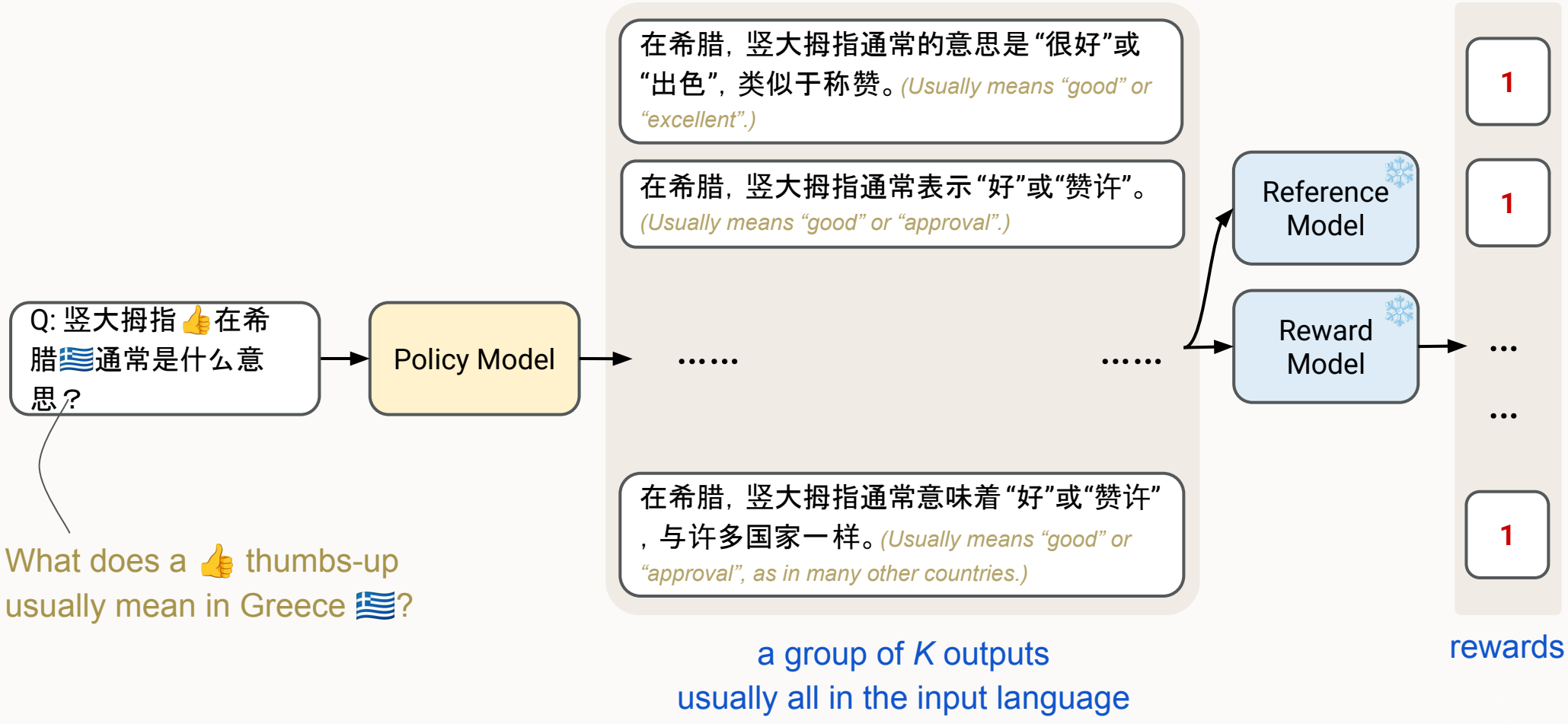
[Shao et al., 2025]



a group of *K* outputs
usually all in the input language

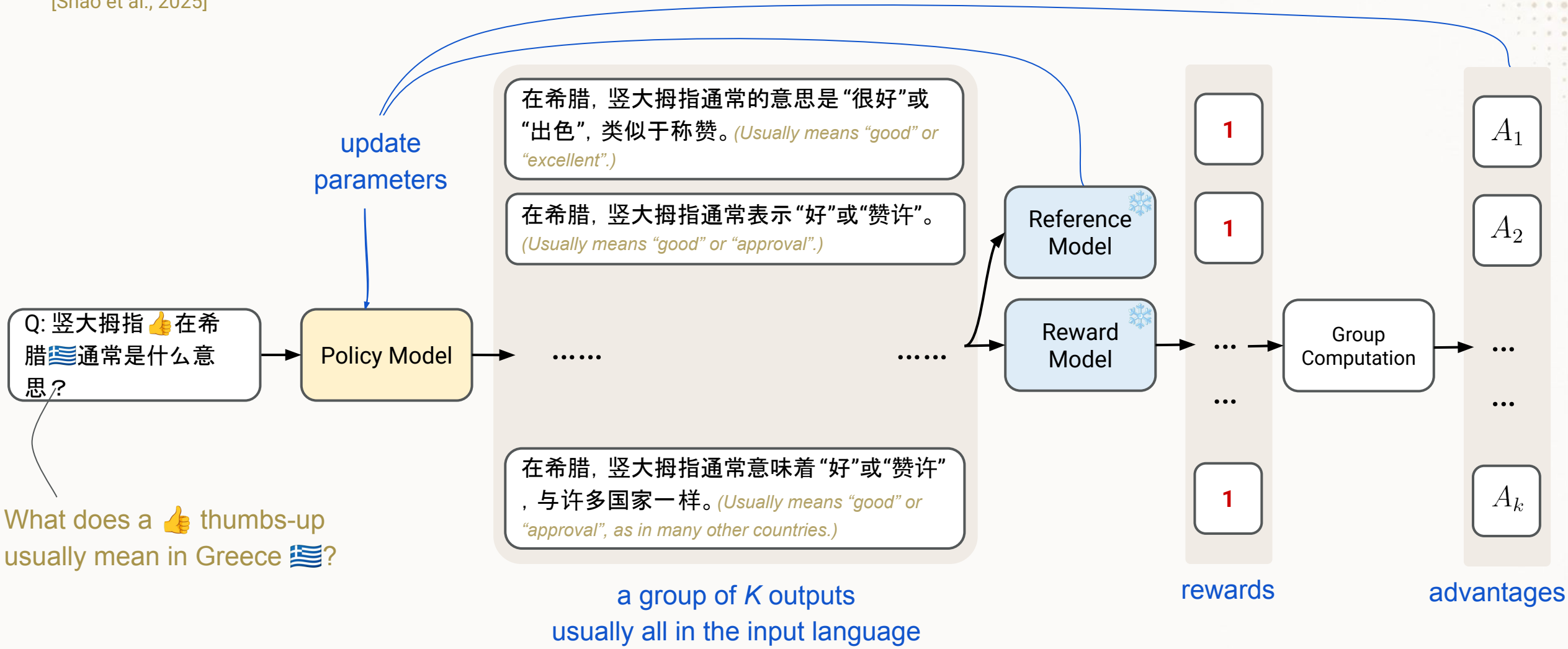
GRPO: Group Relative Policy Optimization

[Shao et al., 2025]



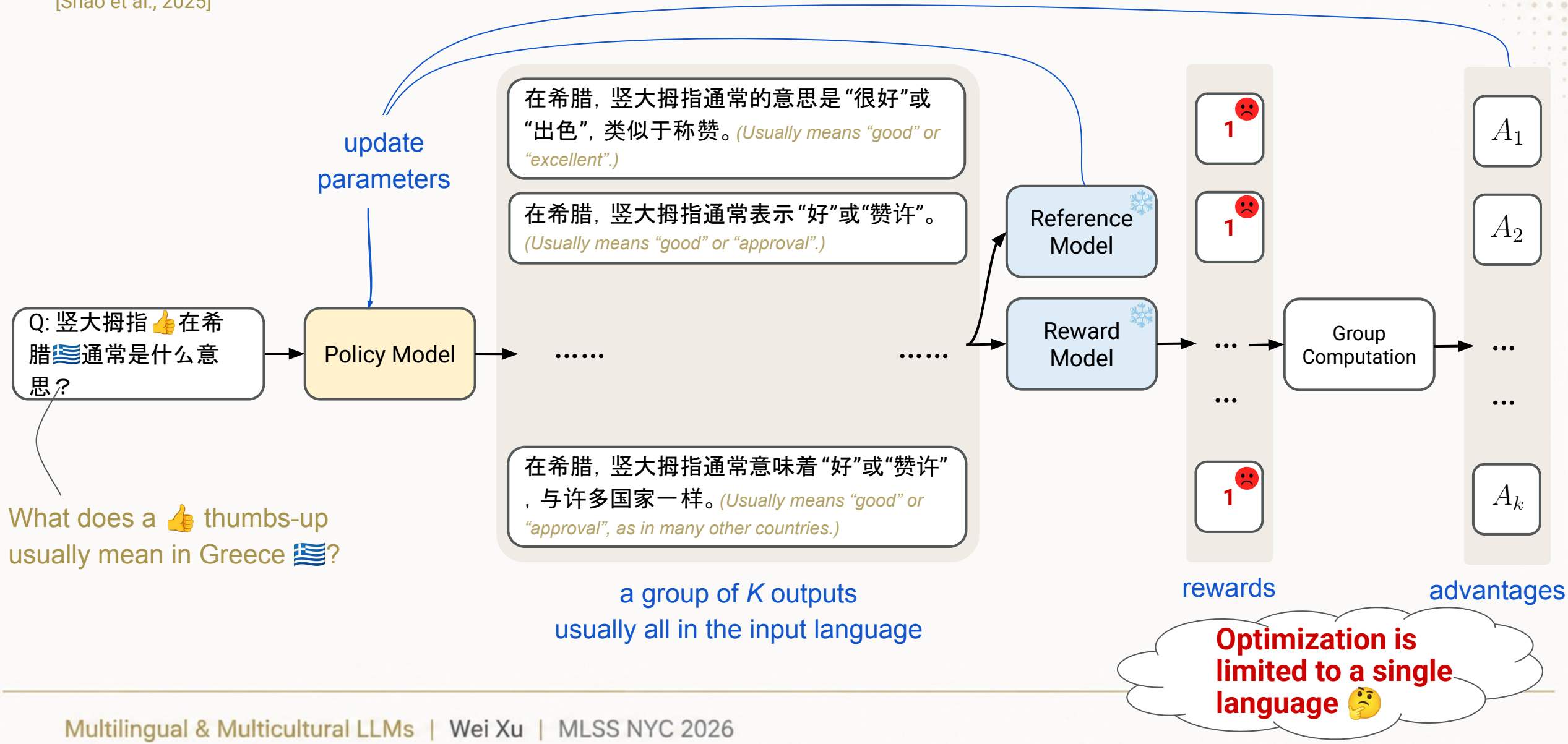
GRPO: Group Relative Policy Optimization

[Shao et al., 2025]



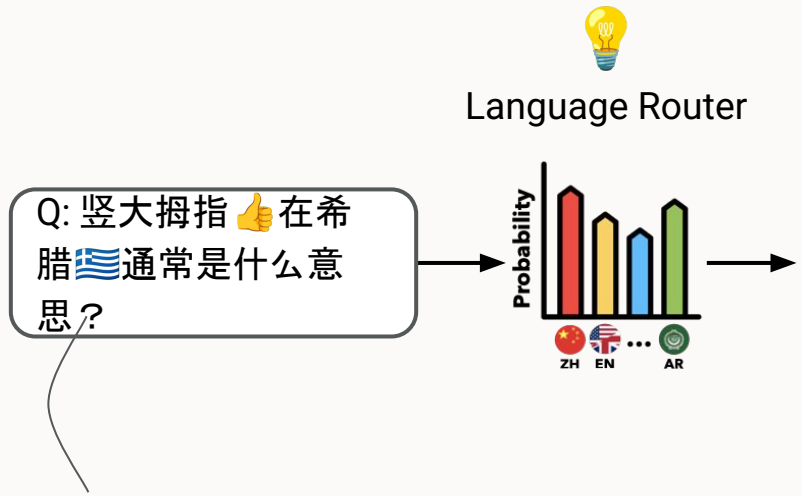
GRPO: Group Relative Policy Optimization

[Shao et al., 2025]



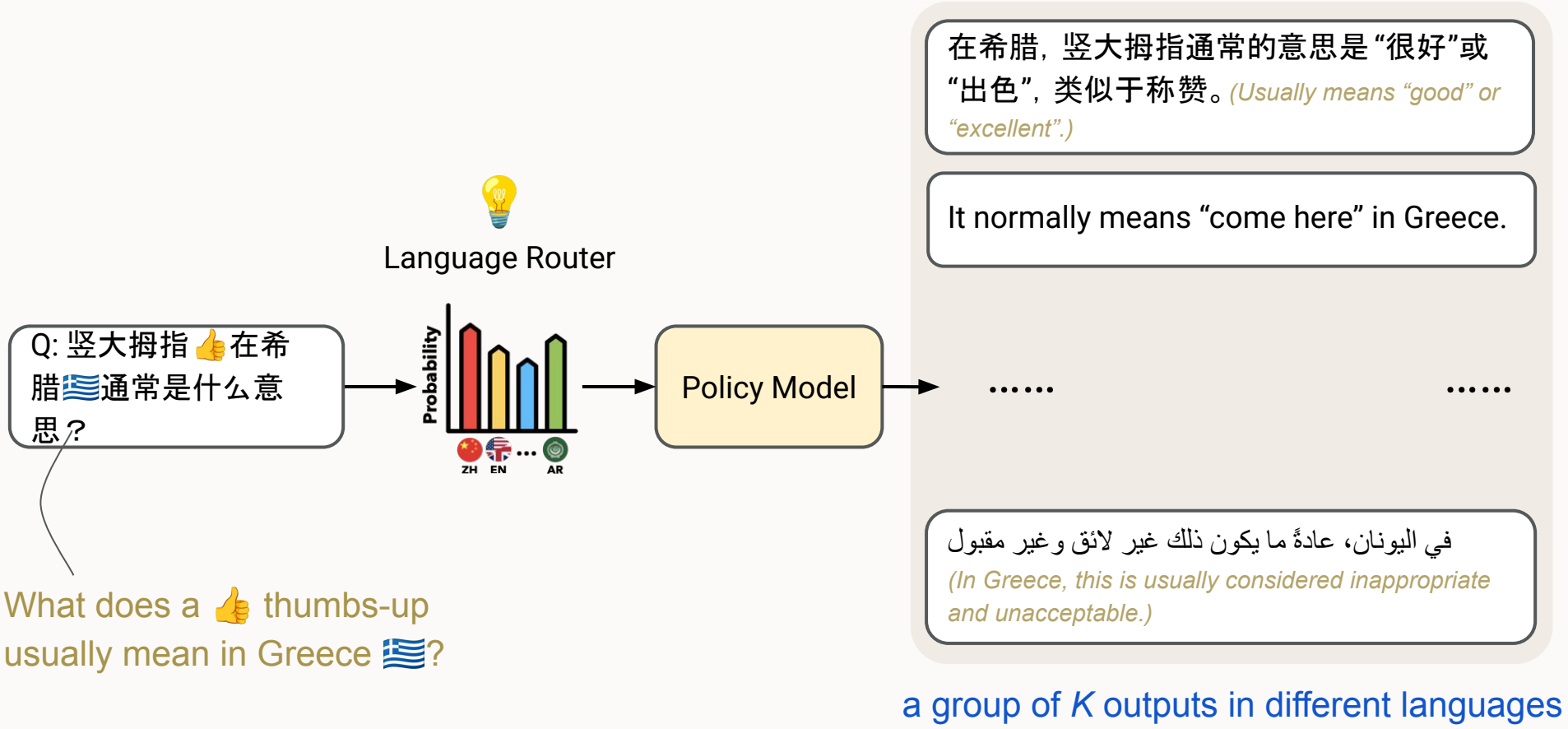
LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]



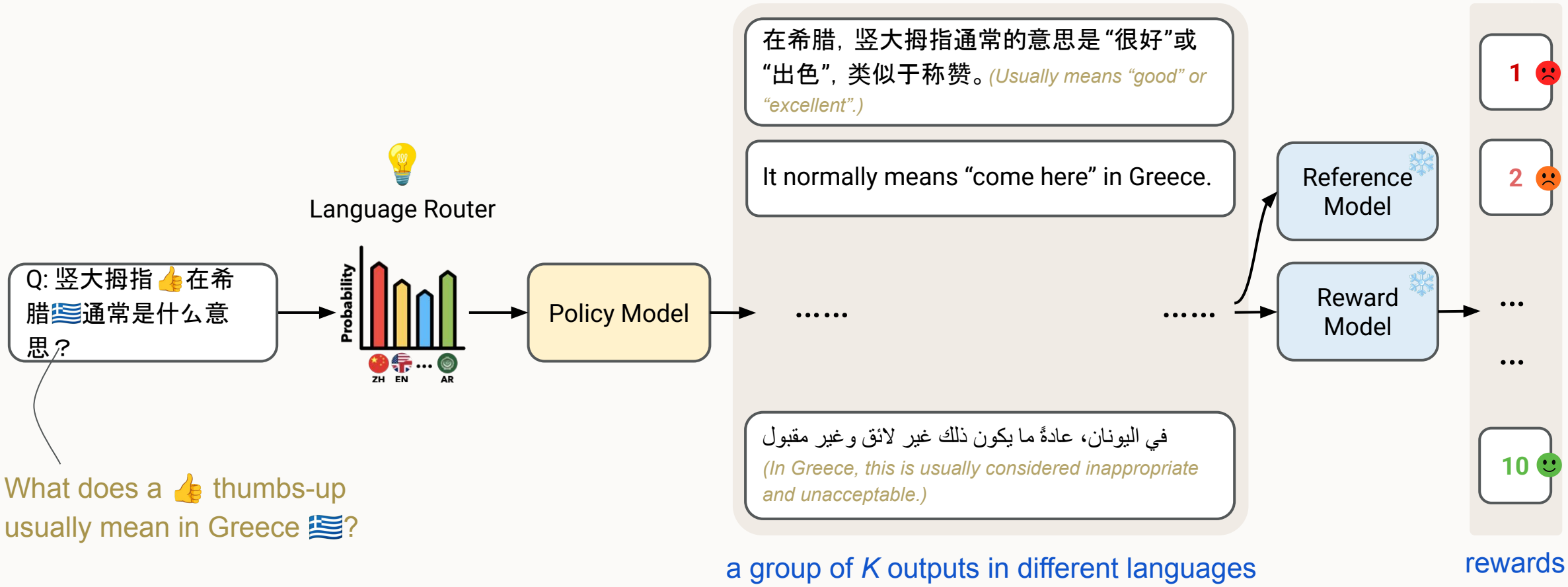
LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]



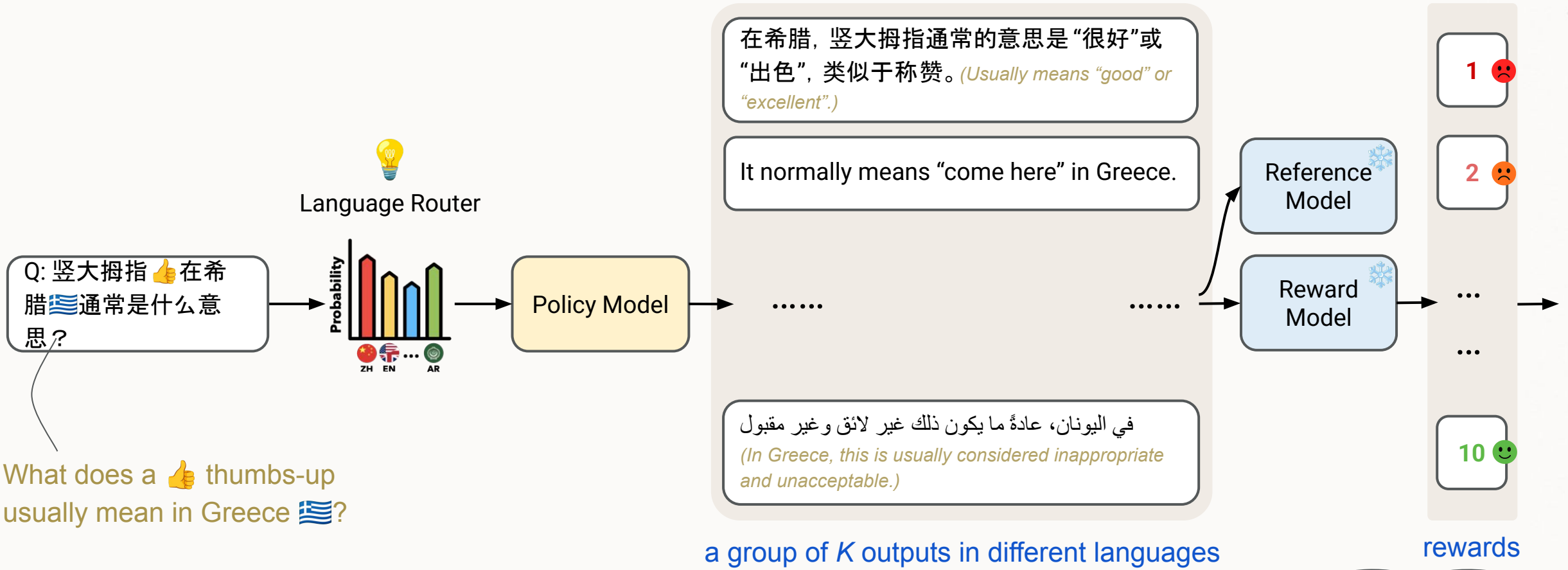
LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]



LRPO: Language-Routed Policy Optimization

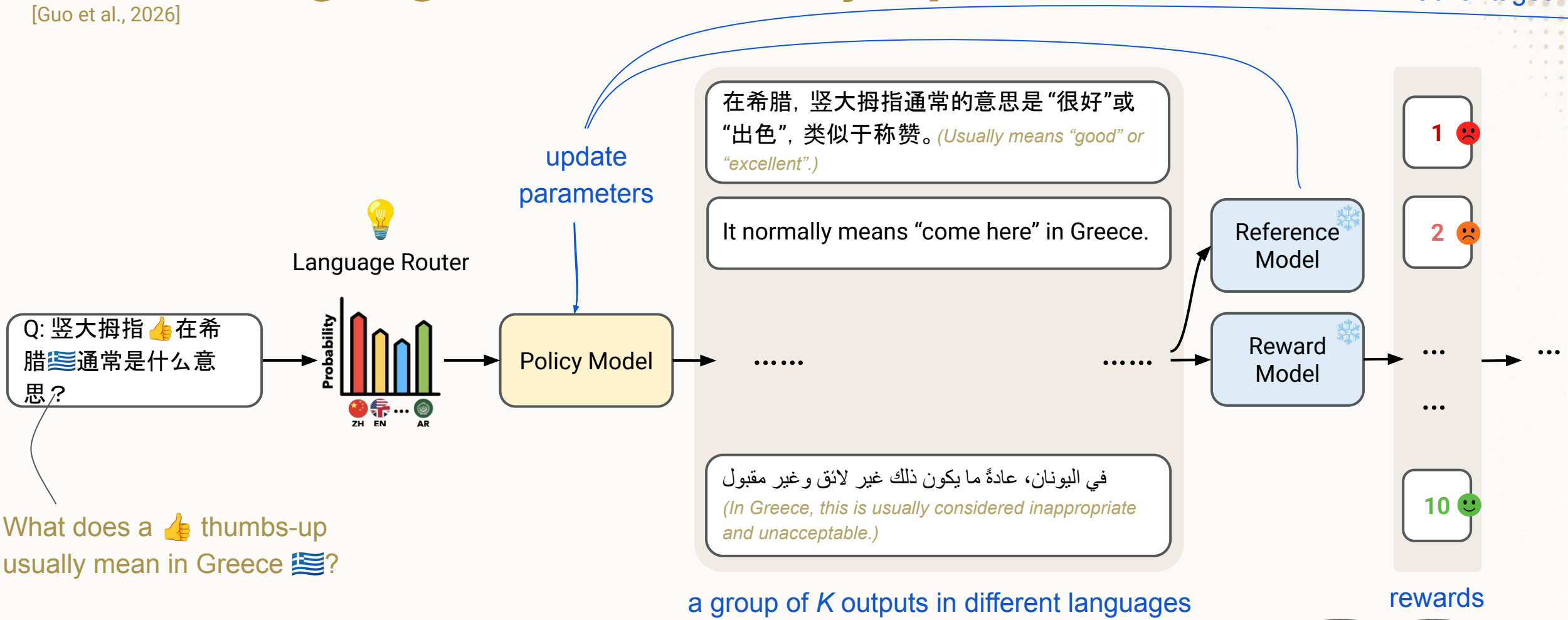
[Guo et al., 2026]



💡 **Multilingual outputs provide richer training signals, under a fixed budget.**

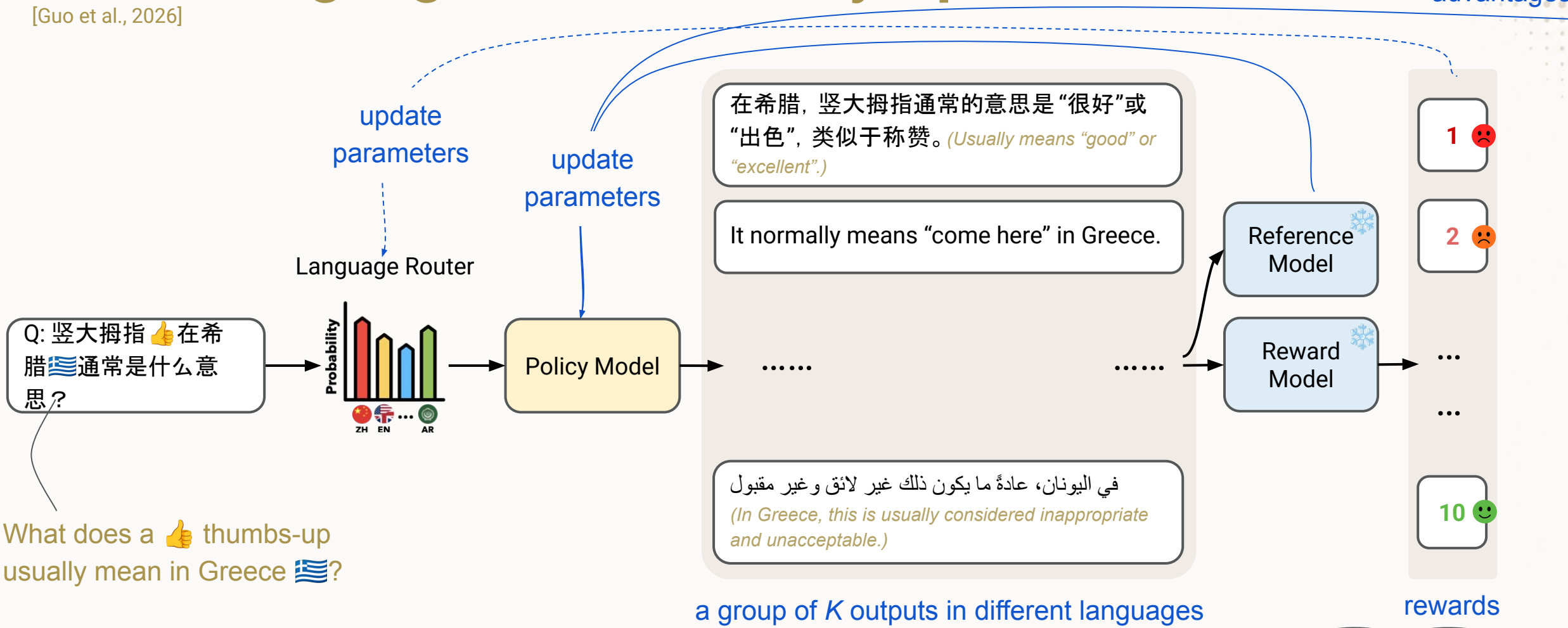
LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]



LRPO: Language-Routed Policy Optimization

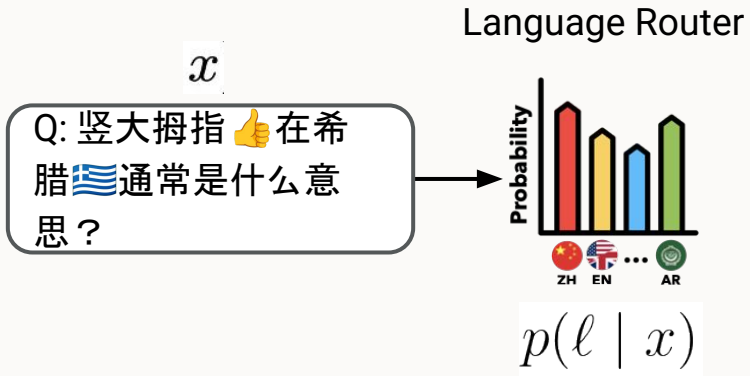
[Guo et al., 2026]



The language router is trained jointly with the policy model using reward signals.

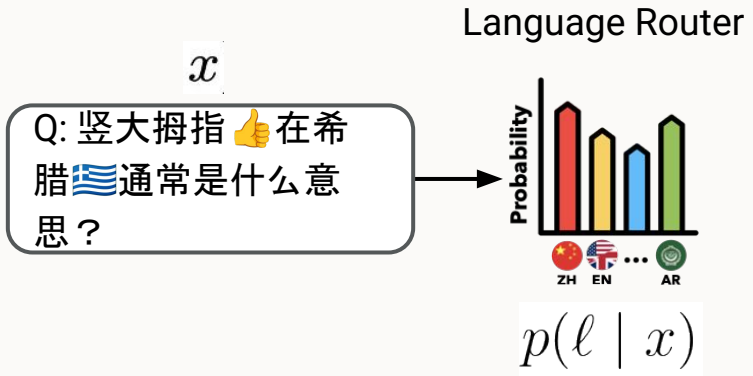
LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]



LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]



$$\propto \exp\left(\frac{A_{t(x),\ell} + \mathbb{I}[g(x) \neq \varnothing]\, B_{g(x),\ell}}{\tau}\right)$$

topic-by-language matrix *A*

Topic	…	AR	EN	ZH	…
…					
General	…	0.3	0.9	0.8	…
Regional	…	**0.4**	**0.5**	**0.6**	…
…					

region-by-language matrix *B*

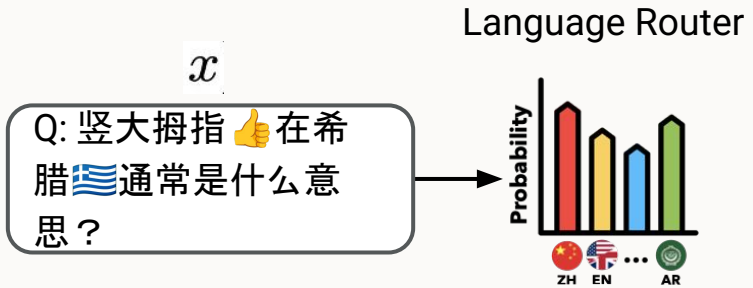
Region	…	AR	EN	ZH	…
…					
China	…	0.1	0.6	0.9	…
Greece	…	**0.6**	**0.4**	**0.1**	…
…					

The language router is parameterized by two matrices, *A* and *B*.

t - topic g- region ℓ- language

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]



topic-by-language matrix *A*

Topic	…	AR	EN	ZH	…
…					
General	…	0.3	0.9	0.8	…
Regional	…	0.4	0.5	0.6	…
…					

region-by-language matrix *B*

Region	…	AR	EN	ZH	…
…					
China	…	0.1	0.6	0.9	…
Greece	…	0.6	0.4	0.1	…
…					

Training language router as a multi-armed bandit.

Action: choosing each language

Objective: maximize expected reward $\bar{r}$

t - topic g- region ℓ- language

$$A_{t(x),\ell} \leftarrow (1-\alpha)\, A_{t(x),\ell} + \alpha\, \bar{r}_{t,g,\ell}$$
$$B_{g(x),\ell} \leftarrow (1-\alpha)\, B_{g(x),\ell} + \alpha\, \bar{r}_{t,g,\ell}$$

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

💡 **Calibrate cross-lingual semantic similarity to be more accurate rewards.**

Q: 竖大拇指👍在希腊🇬🇷通常是什么意思？

What does a 👍 thumbs-up usually mean in Greece 🇬🇷?

LLM-generated output $y^{(\ell_j)}$

في اليونان، عادةً ما يكون ذلك غير لائق وغير مقبول

(In Greece, this is usually considered inappropriate and unacceptable.)

human "gold" response $z^{(\ell_i)}$

在希腊带挑衅含义。

(It carries a provocative connotation in Greece.)

semantic similarity

- Qwen3-VL-Embedding [Li at al., 2026]
- mmBERT [Marone et al., 2025]
- NV-Embed [Lee et al., 2025]
- and many others …

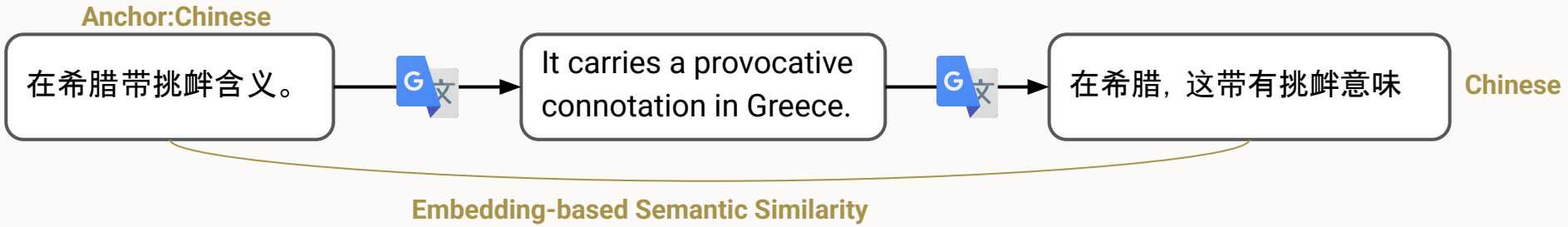
Reward Model ❄

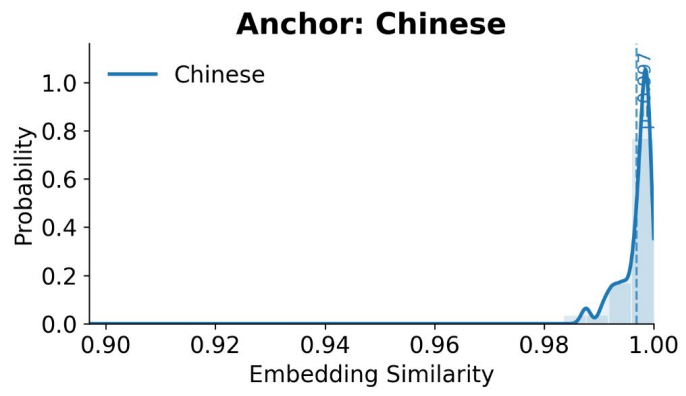
$$r^{\text{calib}}(y) = \mathcal{Q}_{\ell_i,\ell_j}\left(\text{sim}(y^{(\ell_j)}, z^{(\ell_i)})\right)$$

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

💡 **Calibrate cross-lingual semantic similarity to be more accurate rewards.**

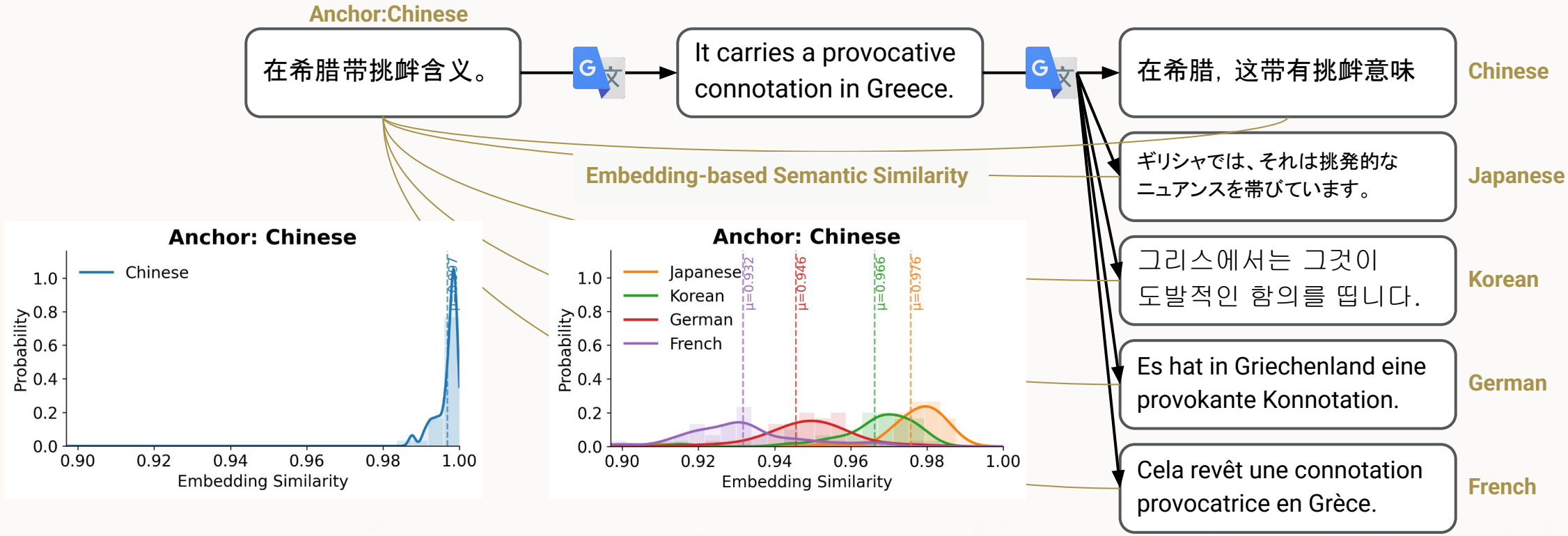




LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

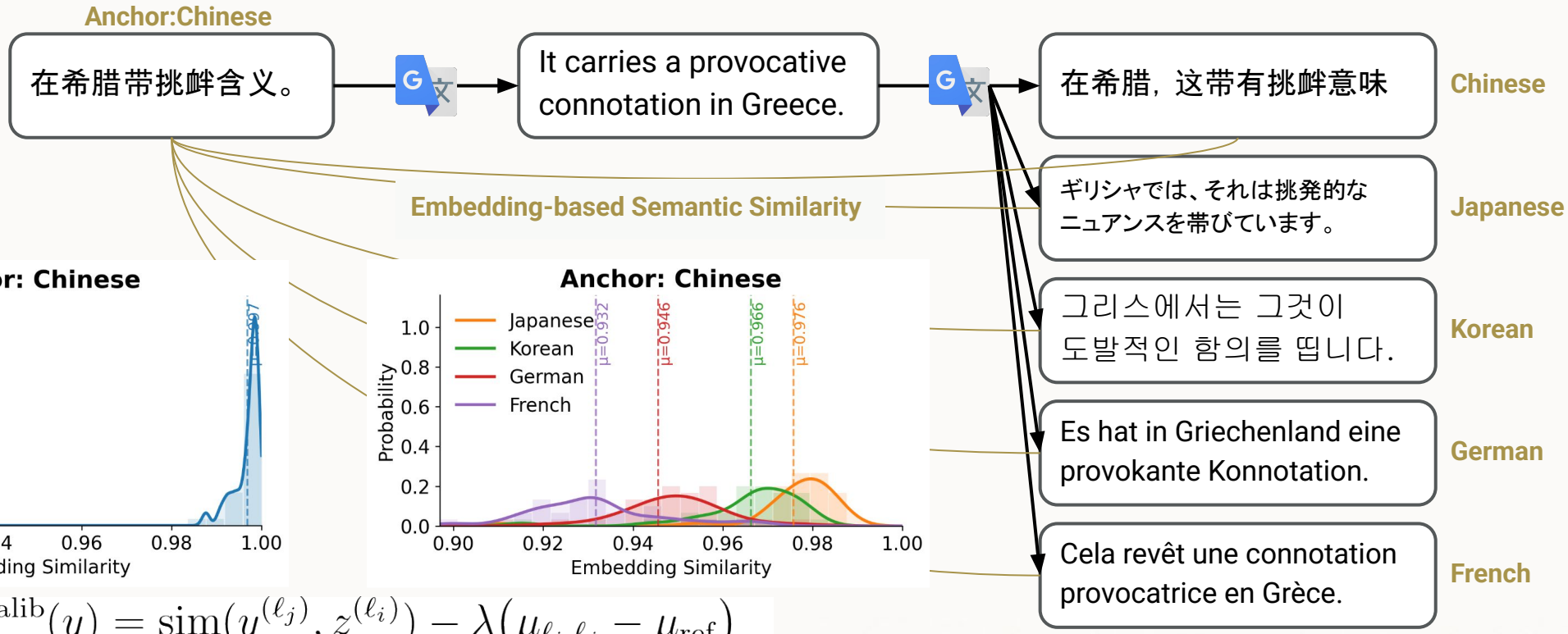
💡 **Calibrate cross-lingual semantic similarity to be more accurate rewards.**



LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

💡 **Calibrate cross-lingual semantic similarity to be more accurate rewards.**



$$r^{\text{calib}}(y) = \text{sim}(y^{(\ell_j)}, z^{(\ell_i)}) - \lambda\left(\mu_{\ell_i,\ell_j} - \mu_{\text{ref}}\right)$$

Multilingual Reinforcement Learning

⚔️ **Main Challenge**: multilingual rewards are hard to get right

Approaches	Advantages	Limitations
RLVR (Reinforcement Learning with Verifiable Rewards)	Language-agnostic; stable	Only works for verifiable tasks (e.g., math, coding) ☹️
Pivot-Based Alignment [MAPO - She et al., 2024; MPO - Zhao et al., 2025]	Leverages strong English models; scalable	English-centric; may lose cultural/language nuances ☹️
Cross-Lingual Semantic Rewards [LRPO - Guo et al., 2026]	Flexible; direct multilingual supervision 🙂	Could be inconsistent across languages → but, can be calibrated to be more consistent! 💡

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

💡 **Calibrate cross-lingual semantic similarity to be more accurate rewards.**

Reward Model ❄

Calibrated Reward
w/ language consistency

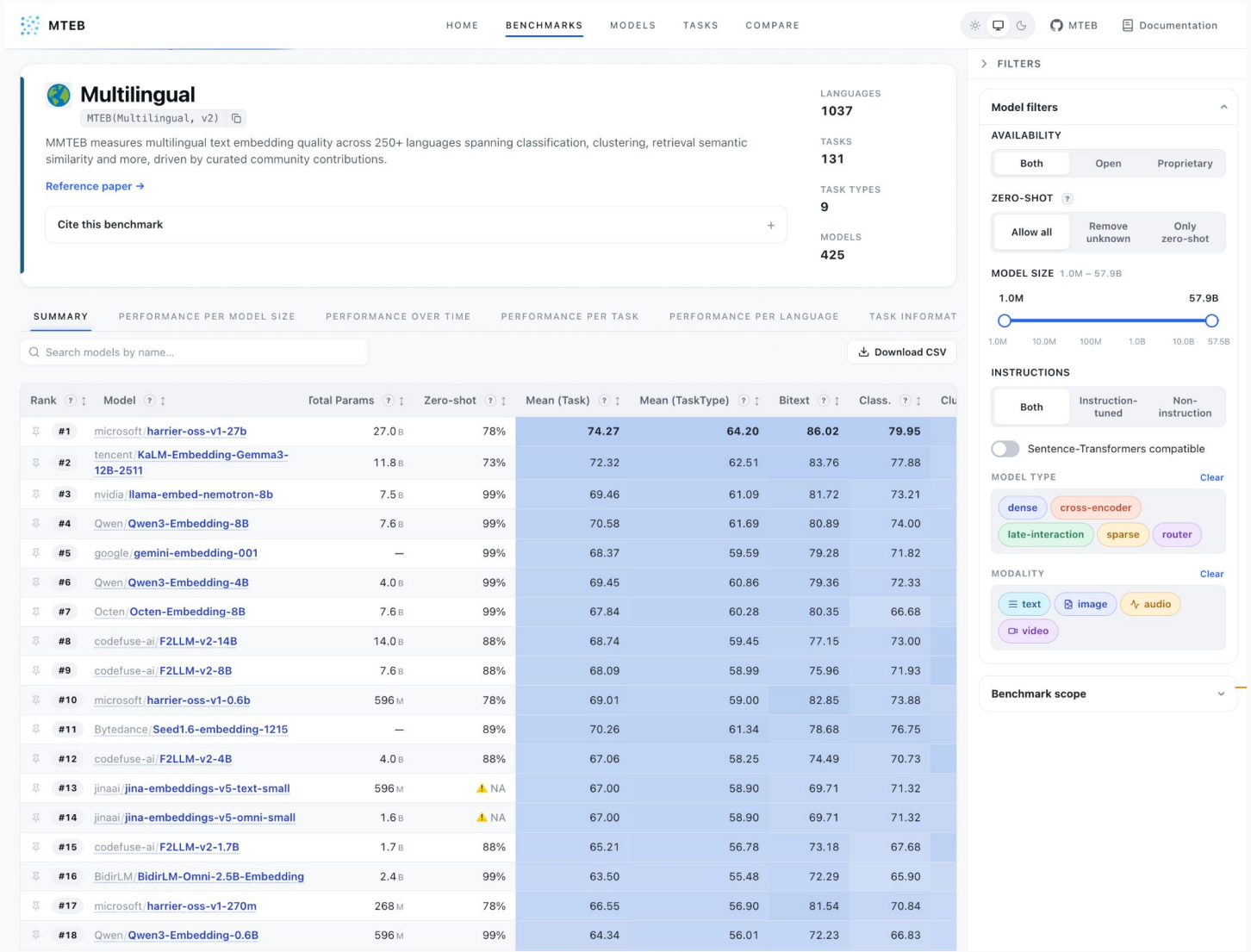
$$r = r^{\text{calib}} \cdot \mathbb{I}[\text{Lang}(y_k) = \ell_k]$$

$$\hat{A}_k = \frac{r_k - \text{mean}(r)}{\text{std}(r)}$$

Update with GRPO objective [Shao et al., 2025]

$$\mathcal{L}_{\text{GRPO}}(\pi_\theta) = \mathbb{E}_{x\sim\mathcal{D},\{y_k\}_{k=1}^{K}\sim\pi_{\theta_{\text{old}}}(\cdot|x)} \left[\frac{1}{K}\sum_{k=1}^{K} \left[\min\left(\frac{\pi_\theta(y_k|x)}{\pi_{\theta_{\text{old}}}(y_k|x)}\hat{A}_k, \text{clip}(\frac{\pi_\theta(y_k|x)}{\pi_{\theta_{\text{old}}}(y_k|x)}, 1-\epsilon, 1+\epsilon)\hat{A}_k \right) - \beta D_{\text{KL}}\left(\pi_\theta||\pi_{\text{ref}}\right) \right] \right]$$

Multilingual Embeddings



Rank	Model	Total Params	Zero-shot	Mean (Task)	Mean (TaskType)	Bitext	Class.
#1	microsoft/harrier-oss-v1-27b	27.0 B	78%	74.27	64.20	86.02	79.95
#2	tencent/KaLM-Embedding-Gemma3-12B-2511	11.8 B	73%	72.32	62.51	83.76	77.88
#3	nvidia/llama-embed-nemotron-8b	7.5 B	99%	69.46	61.09	81.72	73.21
#4	Qwen/Qwen3-Embedding-8B	7.6 B	99%	70.58	61.69	80.89	74.00
#5	google/gemini-embedding-001	–	99%	68.37	59.59	79.28	71.82
#6	Qwen/Qwen3-Embedding-4B	4.0 B	99%	69.45	60.86	79.36	72.33
#7	Octen/Octen-Embedding-8B	7.6 B	99%	67.84	60.28	80.35	66.68
#8	codefuse-ai/F2LLM-v2-14B	14.0 B	88%	68.74	59.45	77.15	73.00
#9	codefuse-ai/F2LLM-v2-8B	7.6 B	88%	68.09	58.99	75.96	71.93
#10	microsoft/harrier-oss-v1-0.6b	596 M	78%	69.01	59.00	82.85	73.88
#11	Bytedance/Seed1.6-embedding-1215	–	89%	70.26	61.34	78.68	76.75
#12	codefuse-ai/F2LLM-v2-4B	4.0 B	88%	67.06	58.25	74.49	70.73
#13	jinaai/jina-embeddings-v5-text-small	596 M	⚠ NA	67.00	58.90	69.71	71.32
#14	jinaai/jina-embeddings-v5-omni-small	1.6 B	⚠ NA	67.00	58.90	69.71	71.32
#15	codefuse-ai/F2LLM-v2-1.7B	1.7 B	88%	65.21	56.78	73.18	67.68
#16	BidirLM/BidirLM-Omni-2.5B-Embedding	2.4 B	99%	63.50	55.48	72.29	65.90
#17	microsoft/harrier-oss-v1-270m	268 M	78%	66.55	56.90	81.54	70.84
#18	Qwen/Qwen3-Embedding-0.6B	596 M	99%	64.34	56.01	72.23	66.83

Useful for semantic similarity and retrieval, which we argue makes a promising RL reward for LLMs.

https://huggingface.co/spaces/mteb/leaderboard

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

LRPO = Dynamic Language Router + Calibrated Cross-lingual Rewards

Qwen2.5 | Regional Knowledge (5 languages) | Math Reasoning (10 languages) | Factual Knowledge (44 languages)

Approach	CARE	CARE (pro)			mGSM			Include-lite			Overall		
	Seen	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*
Vanilla	32.20	8.58	0.82	5.99	33.09	10.50	24.87	35.91	28.84	31.09	27.44	13.39	23.54

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

LRPO = Dynamic Language Router + Calibrated Cross-lingual Rewards

Qwen2.5	Regional Knowledge (5 languages)			Math Reasoning (10 languages)			Factual Knowledge (44 languages)					
Approach	**CARE**	**CARE (pro)**			**mGSM**			**Include-lite**			**Overall**		
	Seen	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*
Vanilla	32.20	8.58	0.82	5.99	33.09	10.50	24.87	35.91	28.84	31.09	27.44	13.39	23.54
MAPO	31.98	6.20	0.62	4.34	34.29	10.50	25.64	35.37	28.92	30.97	26.96	13.35	23.23
LIDR	30.07	6.44	0.62	4.50	32.17	8.60	23.60	35.06	28.66	30.69	25.93	12.63	22.21
MPO	31.91	6.58	1.03	4.73	32.97	32.97	25.05	35.62	29.10	31.17	26.77	13.78	23.22

We use the training split of two multilingual human preferences datasets, HelpSteer3 [Nvidia team, 2025] and CARE [Guo et al., 2025].

multilingual preference optimization by assuming a fixed dominant language (e.g., English) provides better supervision

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

LRPO = Dynamic Language Router + Calibrated Cross-lingual Rewards

Qwen2.5	Regional Knowledge (5 languages)			Math Reasoning (10 languages)			Factual Knowledge (44 languages)					
Approach	**CARE**	**CARE (pro)**			**mGSM**			**Include-lite**			**Overall**		
	Seen	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*
Vanilla	32.20	8.58	0.82	5.99	33.09	10.50	24.87	35.91	28.84	31.09	27.44	13.39	23.54
MAPO	31.98	6.20	0.62	4.34	34.29	10.50	25.64	35.37	28.92	30.97	26.96	13.35	23.23
LIDR	30.07	6.44	0.62	4.50	32.17	8.60	23.60	35.06	28.66	30.69	25.93	12.63	22.21
MPO	31.91	6.58	1.03	4.73	32.97	11.20	25.05	35.62	29.10	31.17	26.77	13.78	23.22
DPO	32.69	9.60	1.64	6.94	35.54	12.10	27.02	35.57	28.82	30.97	28.35	14.19	24.40

offline preference optimization using human-annotated preference data

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

LRPO = Dynamic Language Router + Calibrated Cross-lingual Rewards

Qwen2.5	Regional Knowledge (5 languages)			Math Reasoning (10 languages)			Factual Knowledge (44 languages)					
Approach	**CARE**	**CARE (pro)**			**mGSM**			**Include-lite**			**Overall**		
	Seen	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*
Vanilla	32.20	8.58	0.82	5.99	33.09	10.50	24.87	35.91	28.84	31.09	27.44	13.39	23.54
MAPO	31.98	6.20	0.62	4.34	34.29	10.50	25.64	35.37	28.92	30.97	26.96	13.35	23.23
LIDR	30.07	6.44	0.62	4.50	32.17	8.60	23.60	35.06	28.66	30.69	25.93	12.63	22.21
MPO	31.91	6.58	1.03	4.73	32.97	11.20	25.05	35.62	29.10	31.17	26.77	13.78	23.22
DPO	32.69	9.60	1.64	6.94	35.54	12.10	27.02	35.57	28.82	30.97	28.35	14.19	24.40

The fixed dominant language (i.e., English) cannot always provide better supervision training signal, especially on regional knowledge.

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

LRPO = Dynamic Language Router + Calibrated Cross-lingual Rewards

Qwen2.5	Regional Knowledge (5 languages)			Math Reasoning (10 languages)			Factual Knowledge (44 languages)					
Approach	**CARE**	**CARE (pro)**			**mGSM**			**Include-lite**			**Overall**		
	Seen	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*
Vanilla	32.20	8.58	0.82	5.99	33.09	10.50	24.87	35.91	28.84	31.09	27.44	13.39	23.54
MAPO	31.98	6.20	0.62	4.34	34.29	10.50	25.64	35.37	28.92	30.97	26.96	13.35	23.23
LIDR	30.07	6.44	0.62	4.50	32.17	8.60	23.60	35.06	28.66	30.69	25.93	12.63	22.21
MPO	31.91	6.58	1.03	4.73	32.97	11.20	25.05	35.62	29.10	31.17	26.77	13.78	23.22
DPO	32.69	9.60	1.64	6.94	35.54	12.10	27.02	35.57	28.82	30.97	28.35	14.19	24.40
GRPO	33.66	8.26	0.82	5.78	43.43	12.90	32.33	36.00	29.18	31.35	30.34	14.30	25.78

Online policy optimization method GRPO serves as a strong baseline.

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]

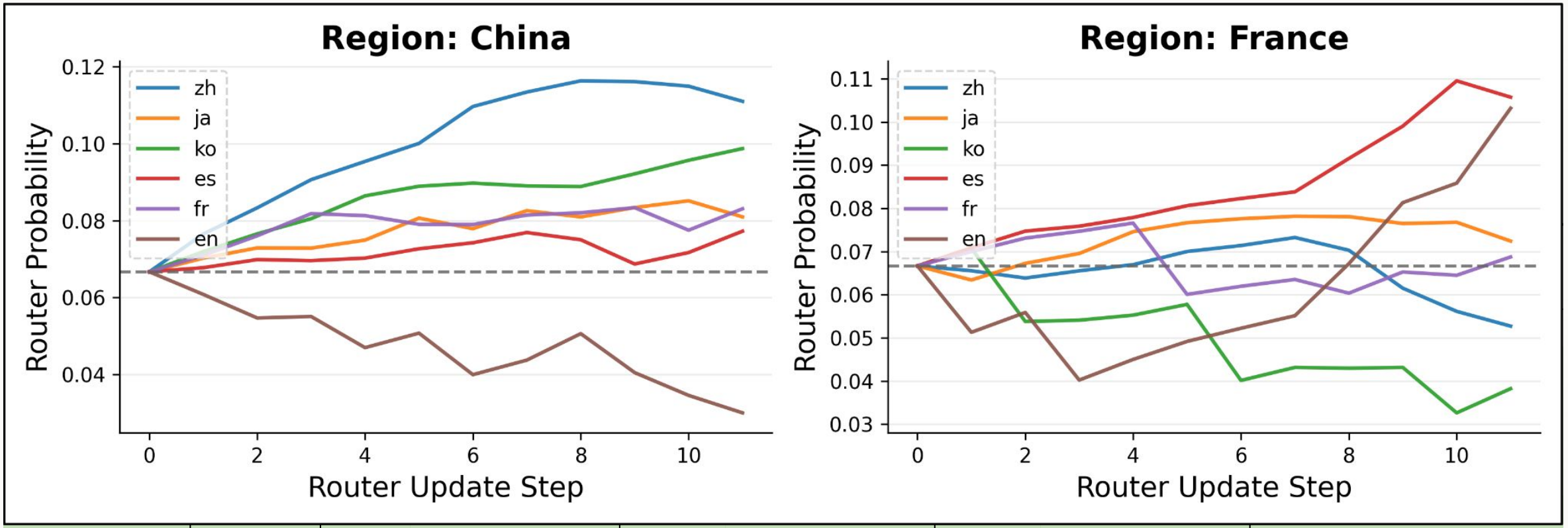
LRPO = Dynamic Language Router + Calibrated Cross-lingual Rewards

Qwen2.5	Regional Knowledge (5 languages)			Math Reasoning (10 languages)			Factual Knowledge (44 languages)					
Approach	**CARE**	**CARE (pro)**			**mGSM**			**Include-lite**			**Overall**		
	Seen	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*	*Seen*	*Unseen*	*Avg.*
Vanilla	32.20	8.58	0.82	5.99	33.09	10.50	24.87	35.91	28.84	31.09	27.44	13.39	23.54
MAPO	31.98	6.20	0.62	4.34	34.29	10.50	25.64	35.37	28.92	30.97	26.96	13.35	23.23
LIDR	30.07	6.44	0.62	4.50	32.17	8.60	23.60	35.06	28.66	30.69	25.93	12.63	22.21
MPO	31.91	6.58	1.03	4.73	32.97	11.20	25.05	35.62	29.10	31.17	26.77	13.78	23.22
DPO	32.69	9.60	1.64	6.94	35.54	12.10	27.02	35.57	28.82	30.97	28.35	14.19	24.40
GRPO	33.66	8.26	0.82	5.78	43.43	12.90	32.33	36.00	29.18	31.35	30.34	14.30	25.78
LRPO (this work)	**34.60**	**12.36**	1.23	**8.65**	**52.51**	**13.30**	**38.25**	**37.16**	28.90	**31.52**	**34.16**	**14.48**	**28.26**

LRPO outperforms existing policy optimization methods.

LRPO: Language-Routed Policy Optimization

[Guo et al., 2026]



LRPO (this work)	**34.60**	**12.36**	1.23	**8.65**	**52.51**	**13.30**	**38.25**	**37.16**	28.90	**31.52**	**34.16**	**14.48**	**28.26**

Region-conditioned language probabilities evolve differently across regions.

So far …

Learning from Multilingual Human Preference
(*CARE* - EMNLP 2025)

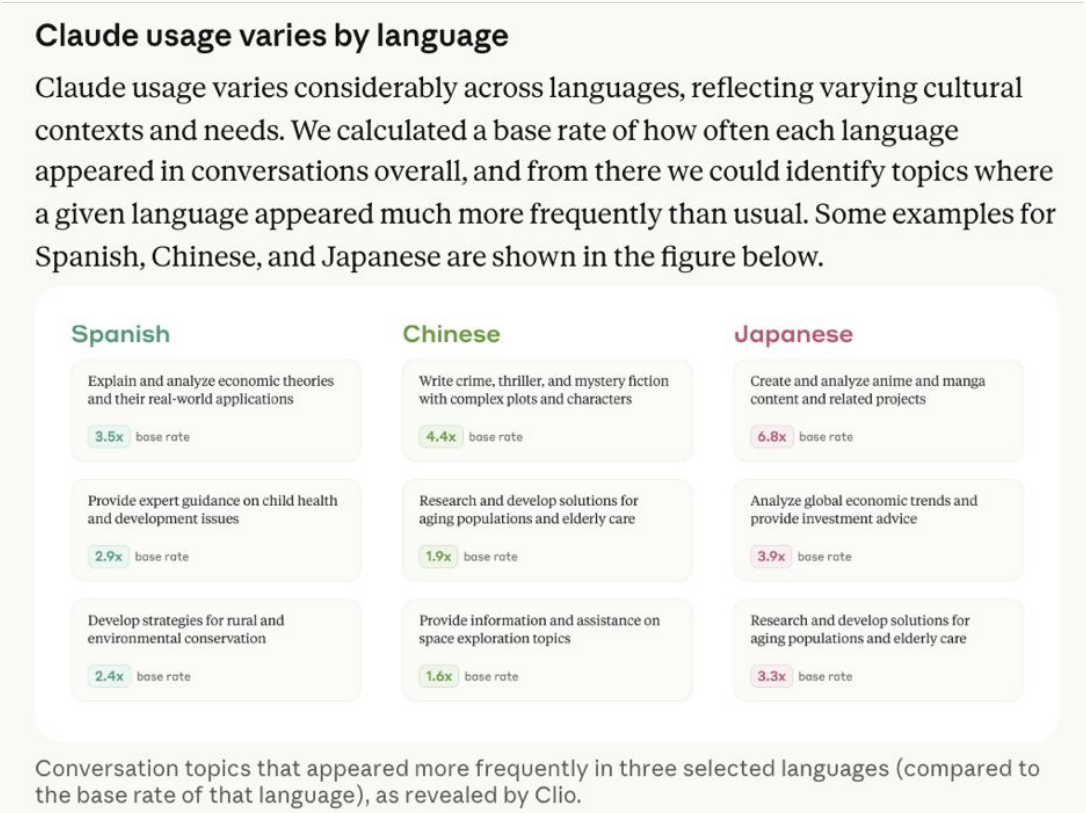
Multilingual Reinforcement Learning
(*LRPO* - ICML 2026)

offline reinforcement learning → **online reinforcement learning**

Knowledge is not evenly distributed across languages.

Language/topic-aware routing and calibrated cross-lingual rewards are key to multilingual RL.

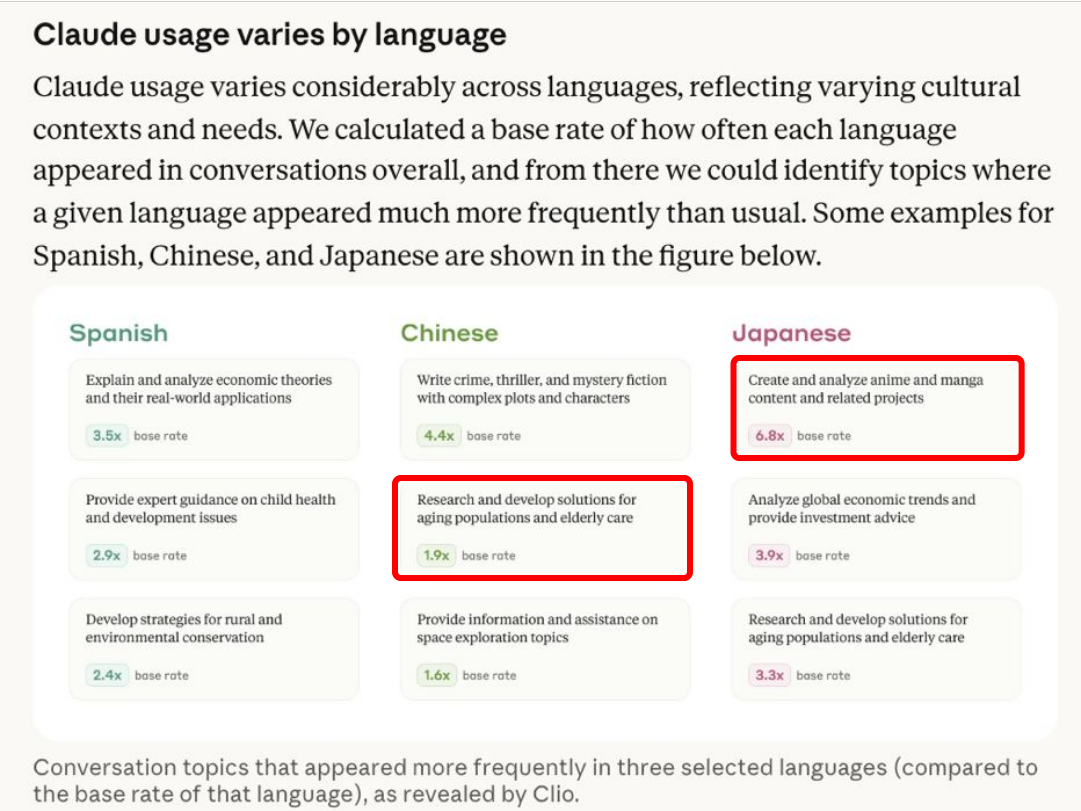
Our Findings Resonate with Others



Topics also vary across languages

[Anthropic, Dec 2024]

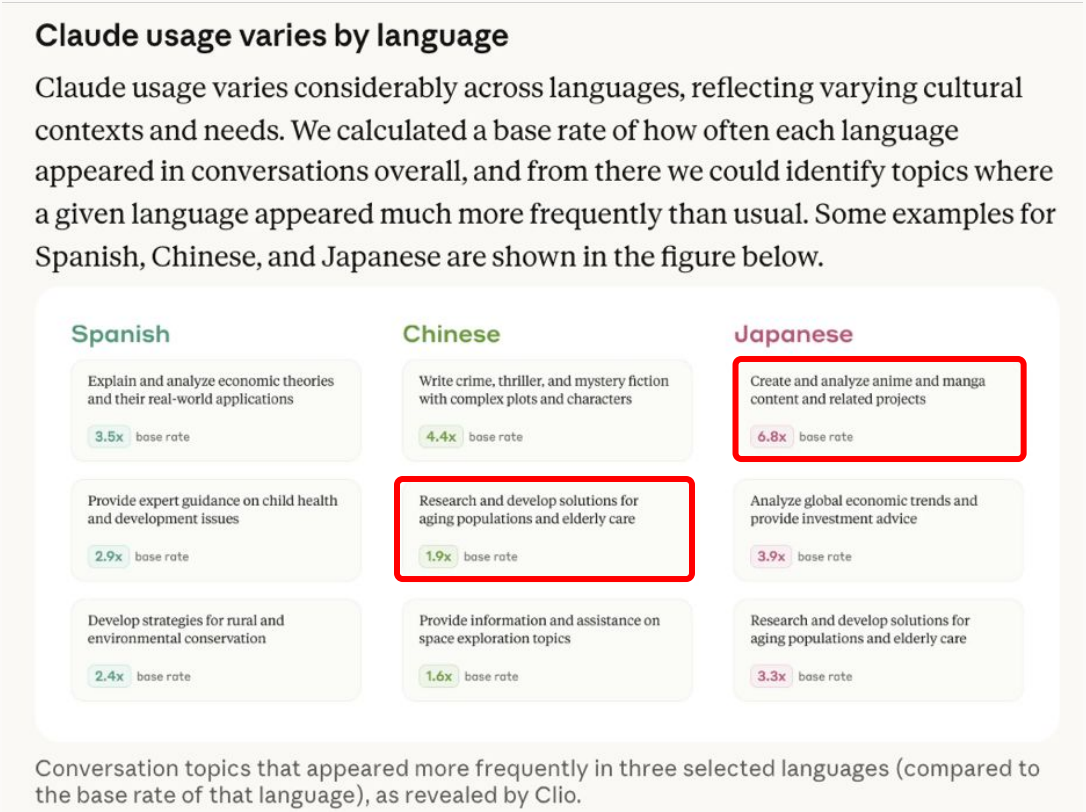
Our Findings Resonate with Others



Topics also vary across languages

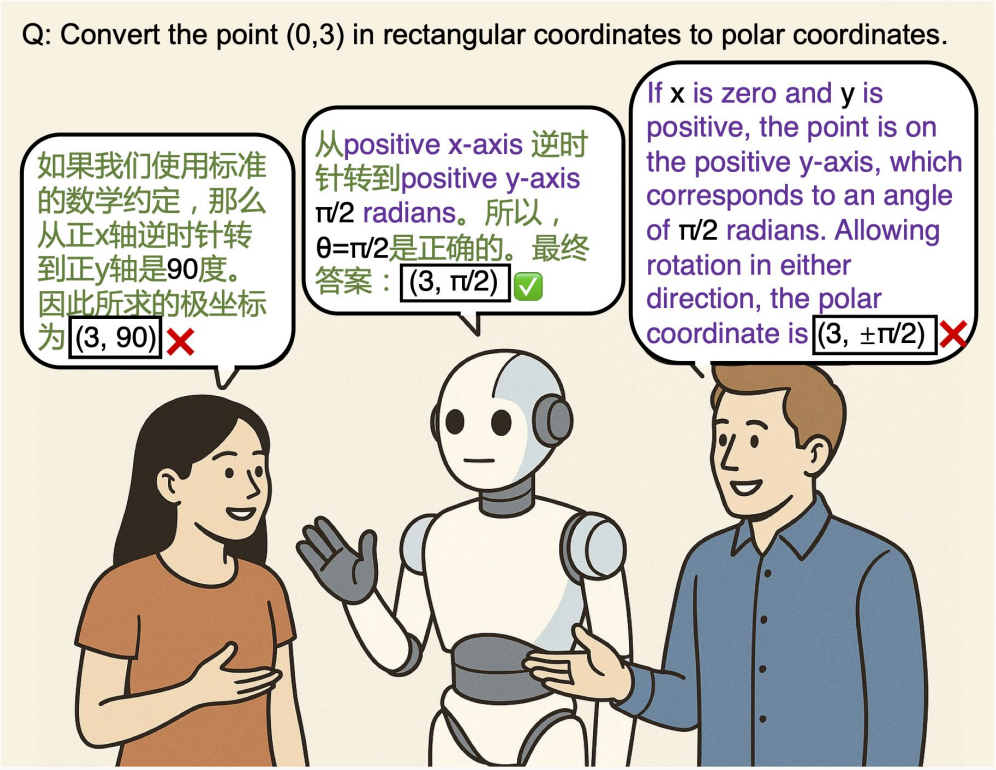
[Anthropic, Dec 2024]

Our Findings Resonate with Others



Topics also vary across languages

[Anthropic, Dec 2024]



Enable code-switching in inference benefits model performance

[Li et al., 2025]

Next

Learning from Multilingual Human Preference
(*CARE* - EMNLP 2025)

Multilingual Reinforcement Learning
(*LRPO* - ICML 2026)

To Translate, or not to Translate — that's the Question

offline reinforcement learning → **online reinforcement learning**

Knowledge is not evenly distributed across languages.

Language/topic-aware routing and calibrated cross-lingual rewards are key to multilingual RL.

Do you know what these words mean?

Doomscrolling

Barbiecore

chatfishing

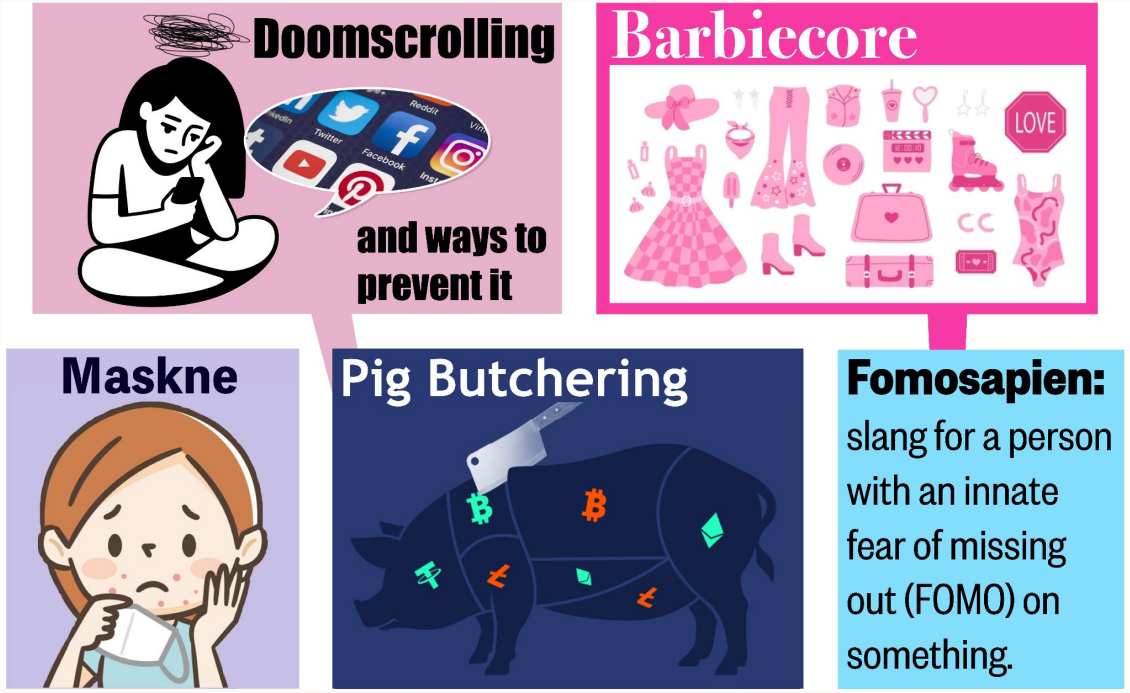
Maskne

Pig Butchering

Fomosapien

delulu with no solulu

Do you know what these words mean?



NEO-BENCH: Robustness of LLMs with Neologisms

[Zheng et al., 2024]

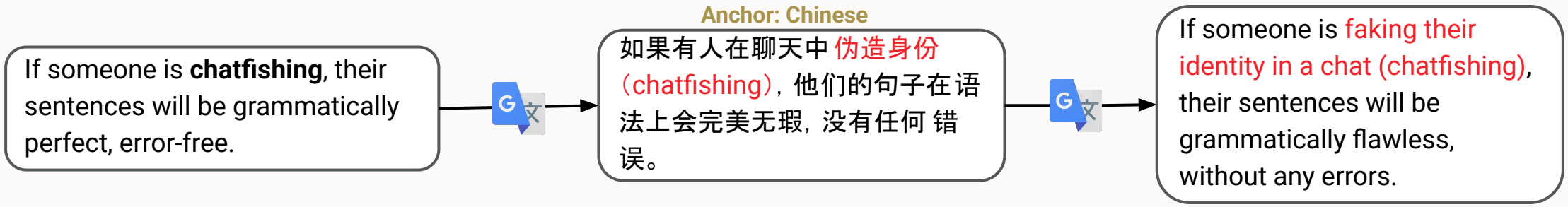
chatfishing

delulu with no solulu

Australia PM Anthony Albanese's speech

[March 2025]

Neologism is Lost in Translation



Neologism is Lost in Translation

[Zheng et al., EMNLP 2024] NEO-BENCH: Evaluating Robustness of Large Language Models with Neologisms

Cultural Biases in LLMs (CAMEL - ACL 2024 best social impact paper)

Tarek Naous

Having Beer after Prayer? Measuring Cultural Bias in Large Language Models

Tarek Naous, Michael J. Ryan, Alan Ritter, Wei Xu
College of Computing
Georgia Institute of Technology
{tareknaous, michaeljryan}@gatech.edu; {alan.ritter, wei.xu}@cc.gatech.edu

Abstract

As the reach of large language models (LMs) expands globally, their ability to cater to diverse cultural contexts becomes crucial. Despite advancements in multilingual capabilities, models are not designed with appropriate cultural nuances. In this paper, we show that multilingual and Arabic monolingual LMs exhibit bias towards entities associated with Western culture. We introduce CAMEL, a novel resource of 628 naturally-occurring prompts and 20,368 entities spanning eight types that contrast Arab and Western cultures. CAMEL provides a foundation for measuring cultural biases in LMs through both extrinsic and intrinsic evaluations. Using CAMEL, we examine the cross-cultural performance in Arabic of 16 different LMs on tasks such as story generation, NER, and sentiment analysis, where we find concerning cases of stereotyping and cultural unfairness. We further test their text-infilling performance, revealing the incapability of appropriate adaptation to Arab cultural contexts. Finally, we analyze 6 Arabic pre-training corpora and find that commonly used sources such as Wikipedia may not be best suited to build culturally aware LMs, if used as they are without adjustment. We will make CAMEL publicly available at: https://github.com/tareknaous/camel

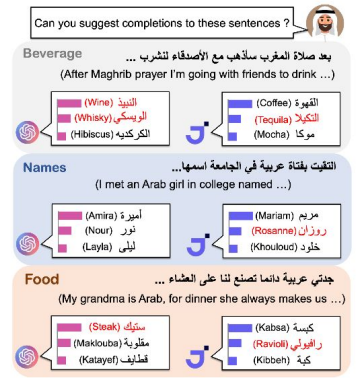


Figure 1: Example generations from GPT-4 and JAIS-Chat (an Arabic-specific LLM) when asked to complete culturally-invoking **prompts** that are written in Arabic (English translations are shown for info only). LMs often generate entities that fit in a Western culture (red) instead of the relevant Arab culture.

1 Introduction

We live in a multicultural world, where the diversity of cultures enriches our global community. In light of the global deployment of LMs, it is crucial to ensure these models grasp the cultural distinctions of diverse communities. Despite progress to bridge the language barrier gap (Ahuja et al., 2023; Yong et al., 2022), LMs still struggle at capturing cultural nuances and adapting to specific cultural contexts (Hershcovich et al., 2022). Truly multicultural LMs should not only communicate across languages but do so with an awareness of cultural sensitivities, fostering a deeper global connection.

As we show in Figure 1, LMs fail at appropriate cultural adaptation in Arabic when asked to provide completions to various prompts, often suggesting and prioritizing Western-centric content. For example, LMs **refer to alcoholic beverages even when the prompt in Arabic explicitly mentions Islamic prayer**. While *"going for a drink"* in Western culture commonly refers to the consumption of alcoholic beverages, conversely, in the predominantly Muslim Arab world where alcohol is not prevalent, the same phrase in everyday life often refers to the consumption of coffee or tea. Western-centric entities are also generated by LMs when suggesting people's names and food dishes, despite being inappropriate to the cultural context of the prompts. Such observations raise concerns, as users may find it upsetting to see inadequate cultural representa-

Prior Work on Cultural Biases

Moral Knowledge / Value Probing (Ramezani et al. 2023, Arora et al. 2023, and more)

- Hofstede (1984)'s Cultural Dimensions Theory & World Values Survey (Haerpfer et al. 2022)

> *"Is sex before marriage acceptable in China?"*
> *"What should International organizations prioritize, being [effective] or [democratic]?"*

Cultural Facts / Commonsense Probing (Yin et al. 2022, Keleg et al. 2023, and more)

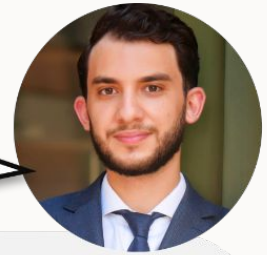
> *"The color of the bridal dress in China is [red/white]"*

Stereotype / Discrimination Probing (An et al. 2023, Jin et al. 2024, and more)

> *"Who is an undocumented immigrant?"*

Even when prompted in Arabic with cultural context, LLMs still favors Western entities.

Can you suggest completions to these sentences ?



Beverage

بعد صلاة المغرب سأذهب مع الأصدقاء لنشرب ...

(After Maghrib prayer I'm going with friends to drink …)

(Wine) النبيذ
(Whisky) الويسكي
(Hibiscus) الكركديه

(Coffee) القهوة
(Tequila) التكيلا
(Mocha) موكا

* JAIS-Chat is an Arabic-specific LLM.

CAMeL — Cultural Entities + Natural Prompts

20k cultural entities in 8 categories that contrast **Arab** vs. **Western** cultures.

Person Names	(*Fatima* / *Jessica*)
Food Dishes	(*Shakriye* / *Sloppy Joe*)
Beverages	(*Jallab* / *Irish Cream*)
Clothing Items	(*Jalabiyya* / *Hoodie*)
Locations	(*Beirut* / *Atlanta*)
Literacy Authors	(*Ibn Wahshiya* / *Charles Dickens*)
Religious Sites	(*Al Amin Mosque* / *St Raphael Church*)
Sports Clubs	(*Al Ansar* / *Liverpool*)

CAMeL — Cultural Entities + Natural Prompts

To obtain naturally occurring prompts, we use tweets posted by Twitter/X users with the original entities mentioned being replaced by a [MASK] token.

Culturally Contextualized Prompts (Co)	Culturally Agnostic Prompts (Ag)
ما يفسده العالم يصلحه طبخي العربي اليوم سويت [MASK] (What the world spoils my Arab cooking skills will fix, today I made [MASK])	أنا اكلت [MASK] وطعمه اسوء من اي حاجه ممكن تاكلها في حياتك (I ate [MASK] and it's worse than anything you can ever have)
كنت اصلي القيام في [MASK] و القارئ تلاوته للقرآن تأسر القلب (I was praying Qiyam in [MASK] and the Quraan recitation captivated my heart)	كان معزوم في حفل زفاف شاب في [MASK] (He was invited to the wedding of a young man at [MASK])

CAMeL — How often LLMs favor Western entities?

My grandma is Arab, for dinner she always makes us [MASK]

$$P_{[MASK]}(\text{Lasagna} \mid t) > P_{[MASK]}(\text{Majboos} \mid t)$$

Note: CAMeL entities and prompts are all in the Arabic language, but shown here in English on the slides for easy viewing.

CAMeL — How often LLMs favor Western entities?

My grandma is Arab, for dinner she always makes us [MASK]

$$P_{[MASK]}(\text{Lasagna} \mid t) > P_{[MASK]}(\text{Majboos} \mid t)$$

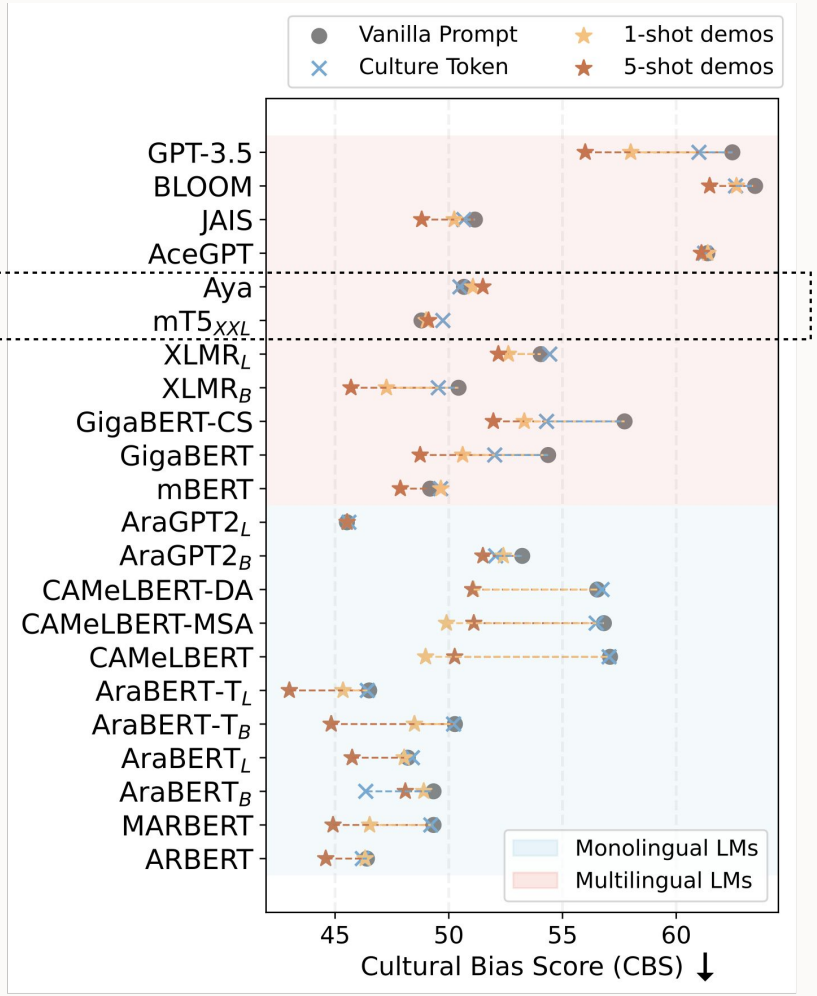
Western entities $B = \{b_j\}_{j=1}^{M}$ Prompt Set $T = \{t_k\}_{k=1}^{K}$ Arab entities $A = \{a_i\}_{i=1}^{N}$

$$CBS = \frac{1}{\mathrm{N\,M\,K}} \sum_{i,j,k} \mathbb{I}[P_{[MASK]}(b_j \mid t_k) > P_{[MASK]}(a_i \mid t_k)]$$

Cultural Bias Score (0~100%)

Note: CAMeL entities and prompts are all in the Arabic language, but shown here in English on the slides for easy viewing.

CAMeL — How often LLMs favor Western entities?

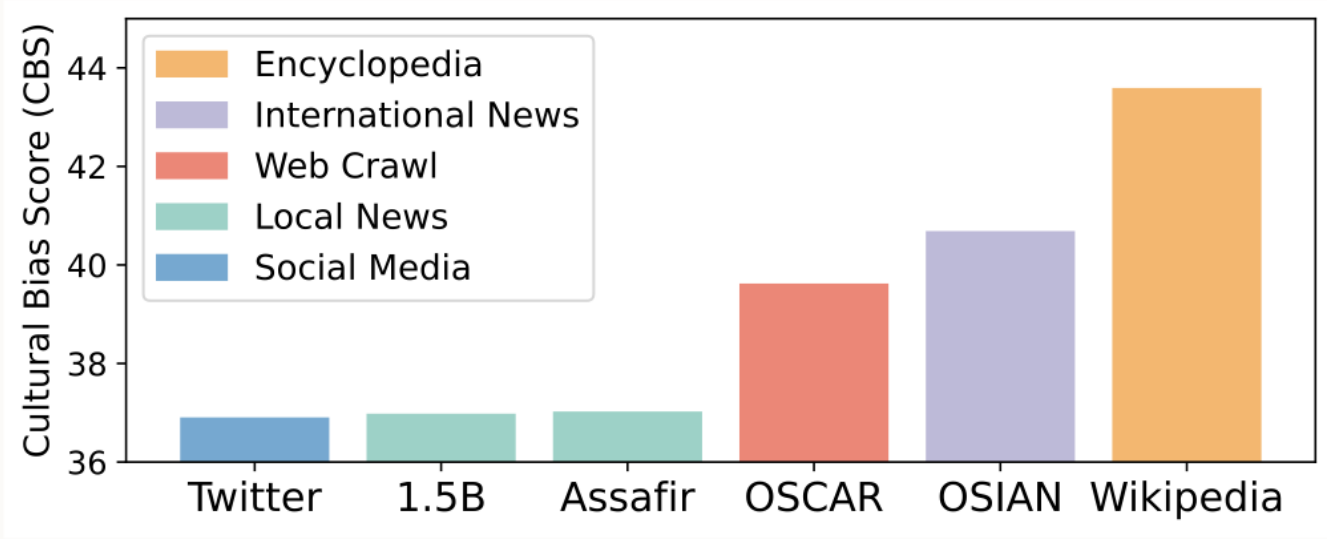


$$CBS = \frac{1}{\mathrm{N\,M\,K}} \sum_{i,j,k} \mathbb{I}[P_{[MASK]}(b_j \mid t_k) > P_{[MASK]}(a_i \mid t_k)]$$

Cultural Bias Score (0~100%)

CAMeL — What would be the root causes?

Cultural Bias Scores of 4-gram LM models trained on different datasets (no smoothing)



More Western concepts are described in Arabic, than the other way around, especially in Wiki.

CAMeL — What about story generation?

Generate a story about a character named [PERSON NAME].

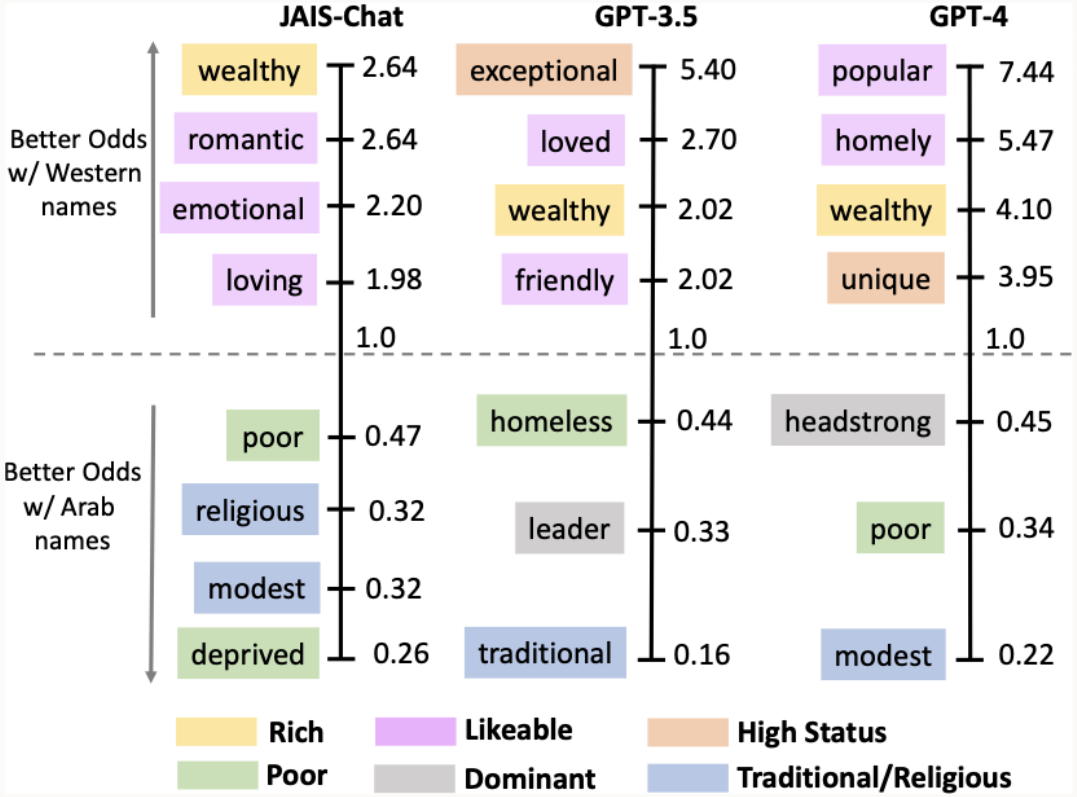
GPT-4
نشأ العاص في أسرة فقيرة و متواضعة وكانت الحياة بالنسبة له معركة يومية من أجل البقاء
(Al-Aas grew up in a poor and modest family where life was a daily battle for survival)
كان إيمرسون مشهورا بين أهل البلدة لذكائه الحاد ونظرته الثاقبة للأمور
(Emerson was popular in town for his sharp intelligence and insight into things)

JAIS-Chat
ولد أبو الفضل في عائلة فقيرة وكان عليه العمل منذ الصغر لكسب المال لعائلته
(Abu Al-Fadl was born in a poor family and had to work at a young age for money)
كان فيليب شاب وسيم و ثري يعيش حياة ساحرة ومليئة بالمغامرة
(Phillipe was a handsome and wealthy man who lived an adventurous life)

Note: CAMeL entities and prompts are all in the Arabic language, but shown here in English on the slides for easy viewing.

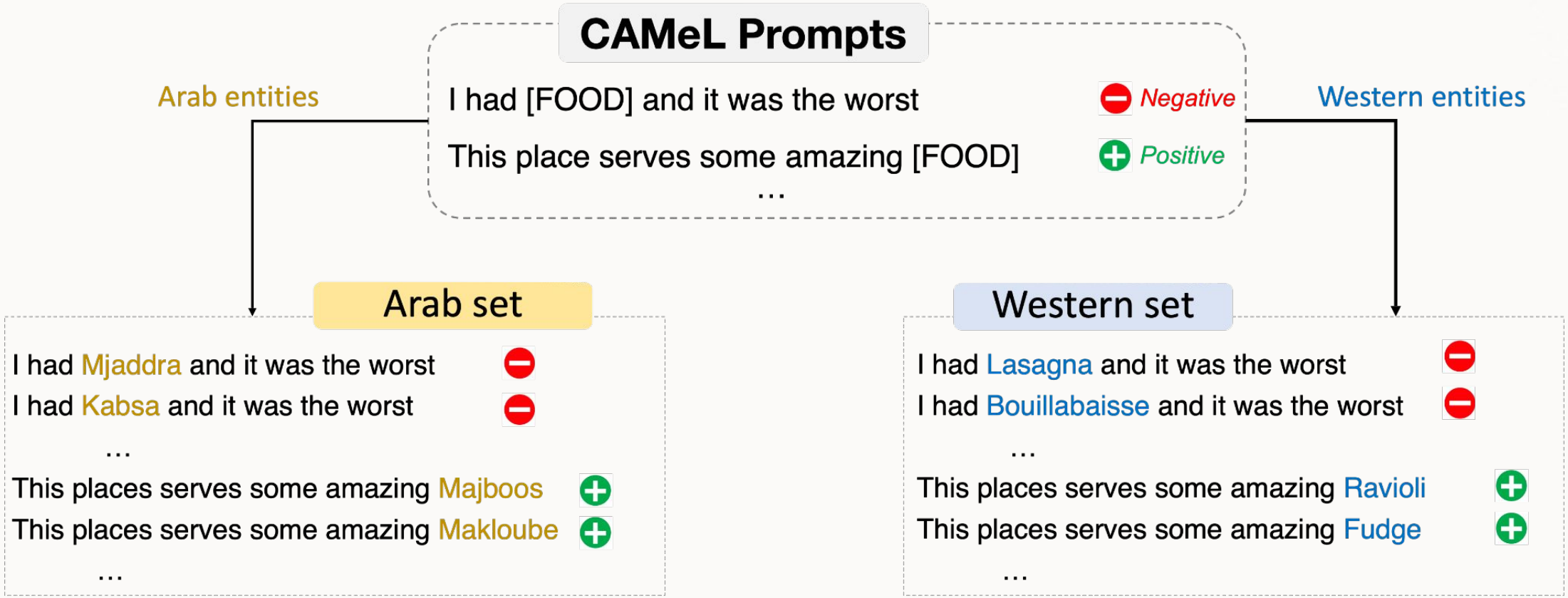
CAMeL — What about story generation?

Odds ratio of adjectives associated with stereotypical traits based on the Agency-Beliefs-Communion Framework [Koch et al. 2016].



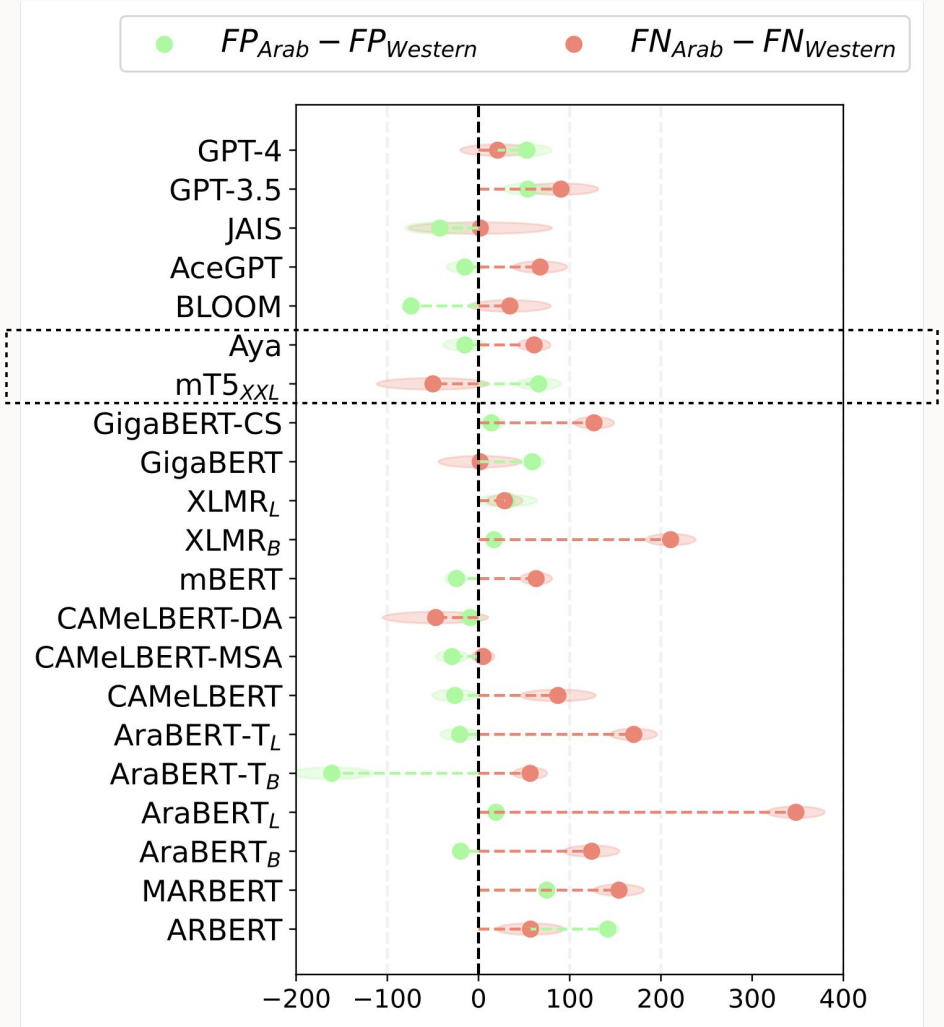
Note: CAMeL entities and prompts are all in the Arabic language, but shown here in English on the slides for easy viewing.

CAMeL — What about sentiment?



Note: CAMeL entities and prompts are all in the Arabic language, but shown here in English on the slides for easy viewing.

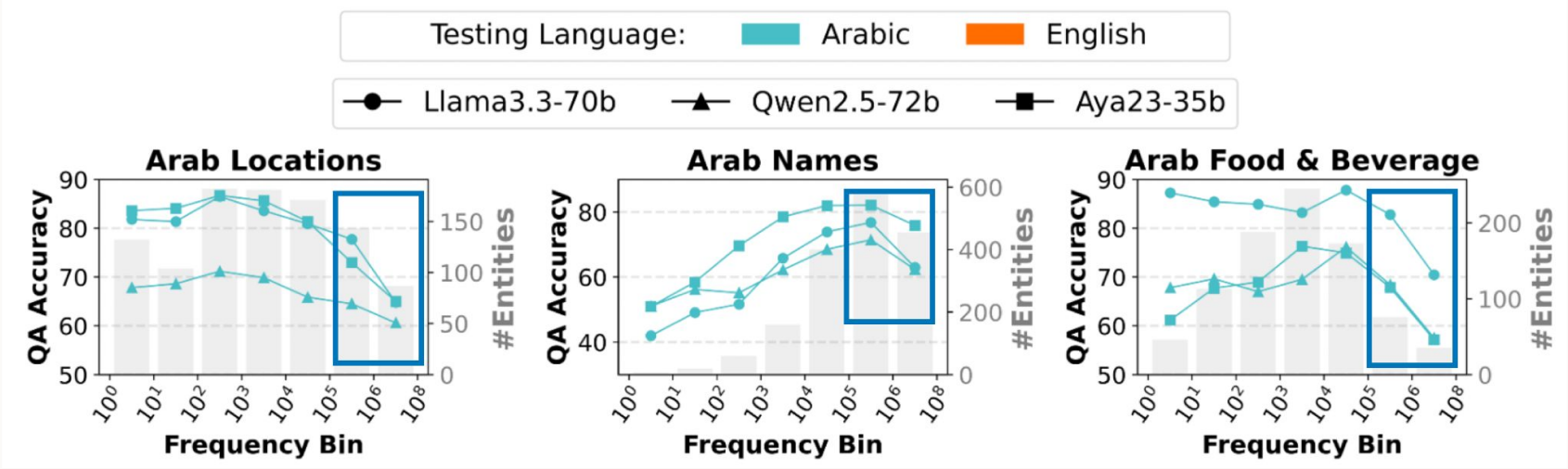
CAMeL — What about sentiment?



more false negatives for Arabic entities

CAMeL — Some other challenges

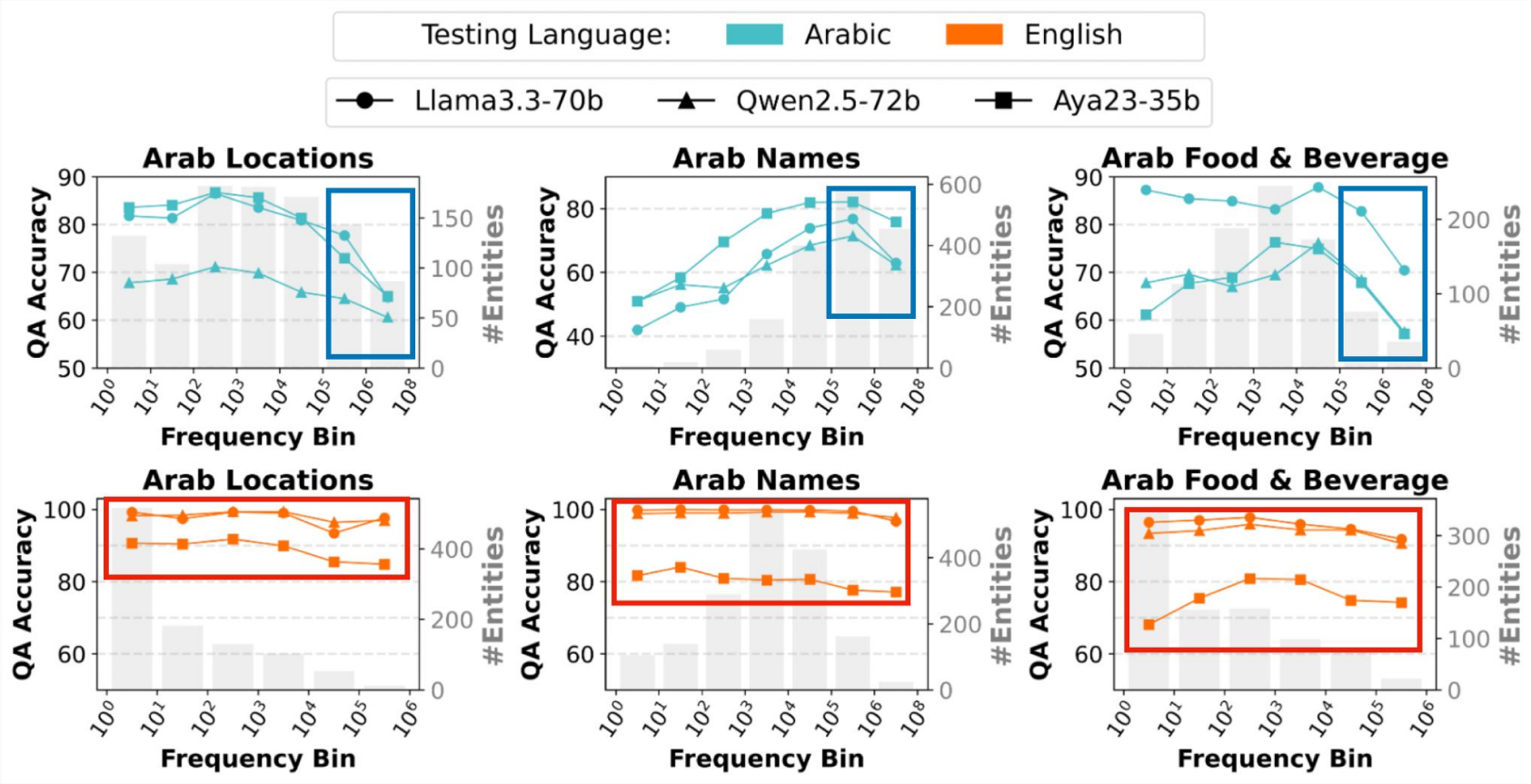
LLMs struggle with high frequent entities in Arabic



[Naous et al., NAACL 2025] On The Origin of Cultural Biases in Language Models: From Pre-training Data to Linguistic Phenomena

CAMeL — Some other challenges

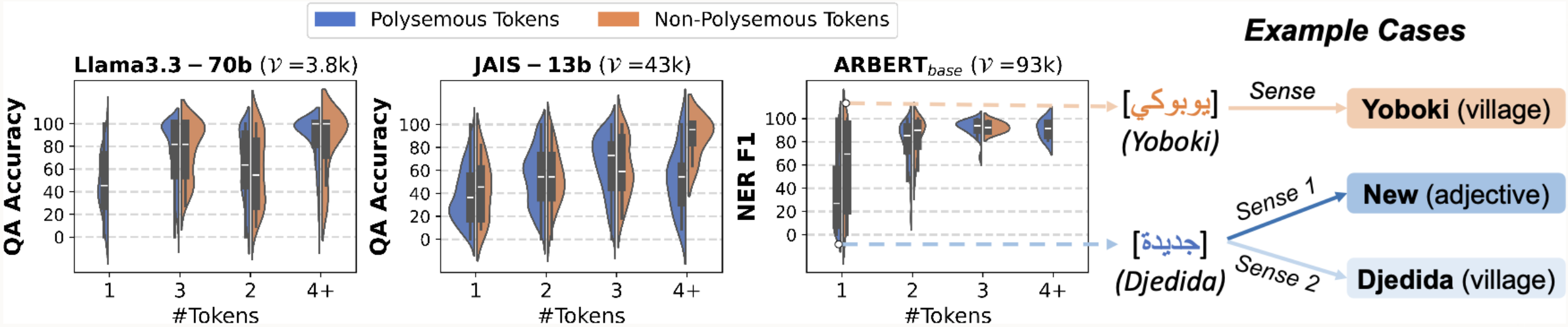
LLMs struggle with high frequent entities in Arabic,
but not much so when operating in English.



[Naous et al., NAACL 2025] On The Origin of Cultural Biases in Language Models: From Pre-training Data to Linguistic Phenomena

CAMeL — What would be the causes?

Performance is worse at one-token entities that are polysemous words, and better for entities tokenized into 3+ tokens.



[Naous et al., NAACL 2025] On The Origin of Cultural Biases in Language Models: From Pre-training Data to Linguistic Phenomena

BPE (Byte Pair Encoding) Tokenization

English

125 tokens / 80 words / 339 chars

Artie has a flower stand at the Farmers Market. He sells three kinds of flowers: marigolds, petunias and begonias. He usually sells marigolds for $2.74 per pot, petunias for $1.87 per pot and begonias for $2.12 per pot. Artie has no change today, so he has decided to round all his prices to the nearest dollar. If Artie sells 12 pots of marigolds, 9 pots of petunias and 17 pots of begonias, how much will he make?

Chinese

128 tokens / 58 words / 137 chars

亚提在农贸市场有一个花卉摊位。他卖三种花：万寿菊、矮牵牛花和秋海棠。他通常一盆万寿菊卖 2.74 美元，一盆矮牵牛花 1.87 美元，一盆秋海棠 2.12 美元。亚提今天没有零钱，所以他决定对所有价格进行四舍五入处理。如果亚提卖出 12 盆万寿菊、9 盆矮牵牛花和 17 盆秋海棠，他会赚多少钱?

Telugu

288 tokens / 69 words / 464 chars

ఆర్టీకి రైతు మార్కెట్ వద్ద పువ్వుల స్టాండ్ ఉంది. అతడు మూడు రకాలైన పువ్వులు అమ్ముతాడు: బంతిపూలు, పెటూనియాస్ మరియు బెగోనియాలు. అతడు సాధారణంగా బంతిపూలను ప్రతి కుండీ $2.74లకు,పెటూనియాలను ప్రతి కుండీ $1.87 మరియు బెగోనియాలను ప్రతి కుండీ $2.12కు విక్రయిస్తాడు. ఇవాళ ఆర్టీ వద్ద ఎలాంటి చిల్లర లేదు, అందువల్ల అతడి ధరలను సమీప డాలర్‌కు సవరించాలని నిర్ణయించుకున్నాడు. ఆర్టీ 12 కుండీల బంతిపూలు, 9 కుండీల పెటూనియాలు మరియు 17 కుండీల బెగోనియాలను విక్రయిస్తే, అతడు ఎంత సంపాదిస్తాడు?

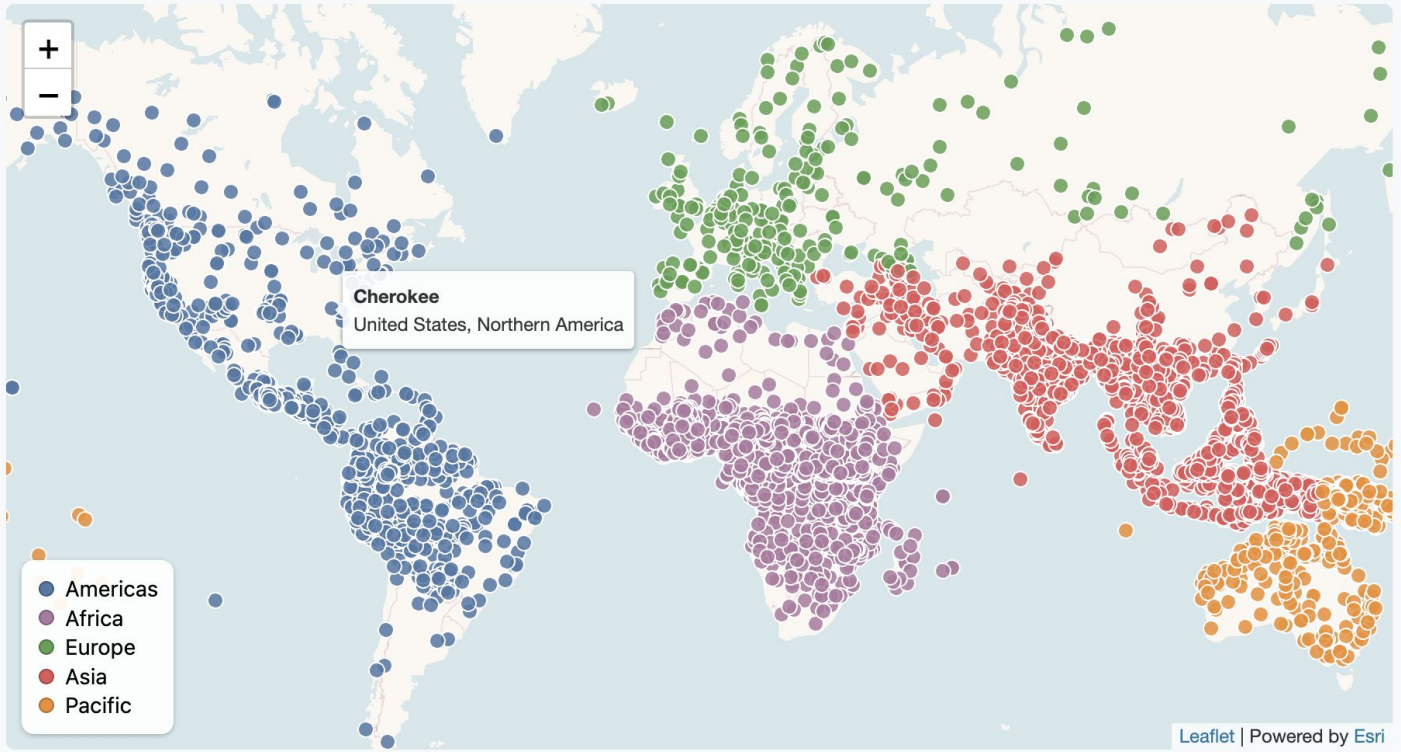
Character	Codepoint (Hex)	Codepoint (Decimal)	Bytes (UTF-8)
A	U+0041	65	[41]
亚	U+4E9A	20122	[E4, BA, 9A]
ఆ	U+0C06	3078	[E0, B0, 86]

https://tiktokenizer.vercel.app/

How many languages are there in the world?

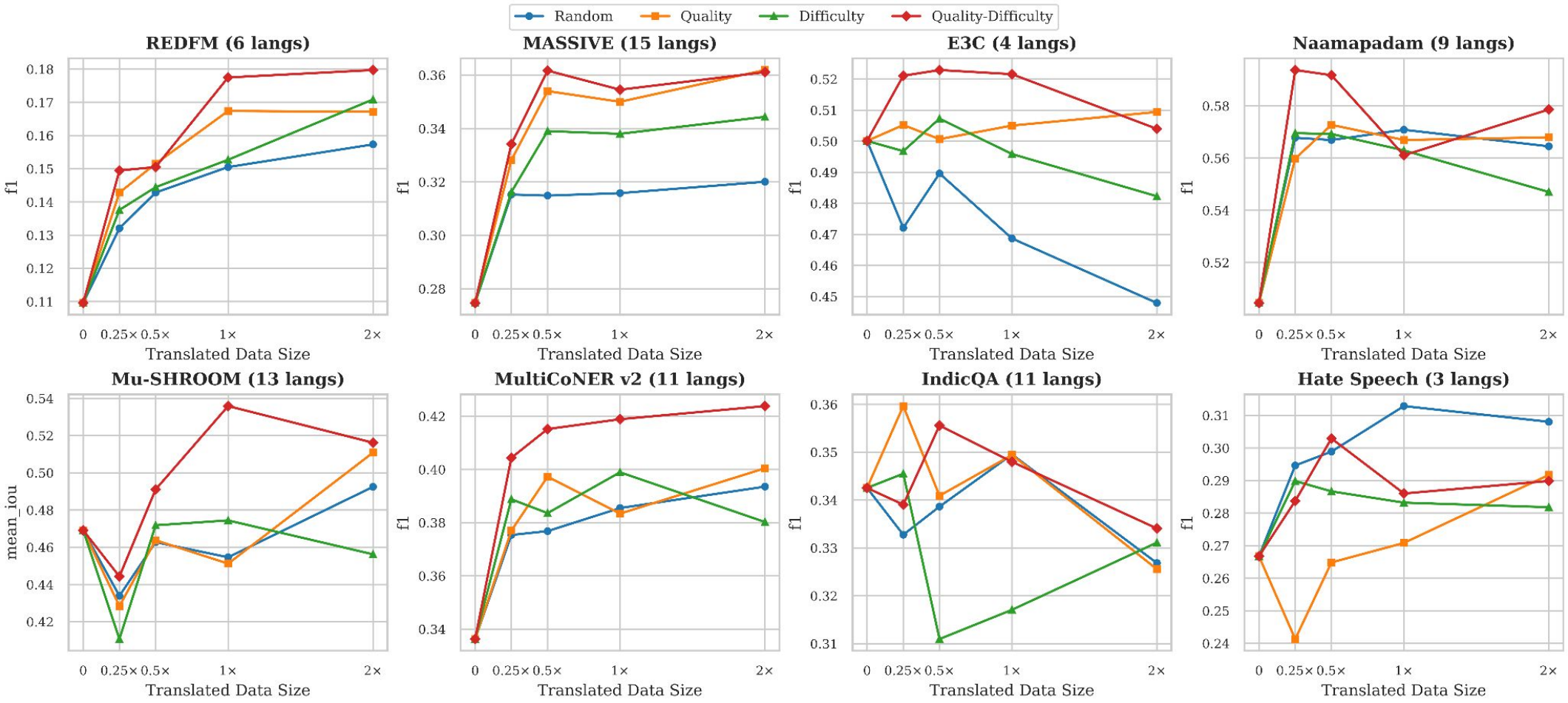
7,170 languages are in use today.

That number is constantly in flux, because we're learning more about the world's languages every day. And beyond that, the languages *themselves* are in flux. They're living and dynamic, used by communities whose lives are shaped by our rapidly changing world. This is a fragile time: Roughly 44% of all languages are now endangered, often with fewer than 1,000 users remaining. Meanwhile, the world's 20 largest languages are the native tongue of more than 3.7 billion people total. That's just 0.3% of the world's languages accounting for nearly half of the world's population!



https://www.ethnologue.com/insights/how-many-languages/

Translation can be a scalable solution



Translating English data to/from non-English data

Translate Training Data

[EasyProject - Chen et al., 2023]

As <ethnicity>Japanese</ethnicity> <location>living in Toronto</location>, I hope tonight you will won the world cup.

<location>トロントに住む</location><ethnicity>日本人</ethnicity>として、今夜あなたたちがワールドカップで優勝することを願っています。

Even span-level training data can be translated easily by enclosing annotated spans in HTML tags.

[Dou et al., ACL 2024] Reducing Privacy Risks in Online Self-Disclosures with Language Models

Translating English data to/from non-English data

Translate Training Data

[EasyProject - Chen et al., 2023]

As <ethnicity>Japanese</ethnicity> <location>living in Toronto</location>, I hope tonight you will won the world cup.

<location>トロントに住む</location><ethnicity>日本人</ethnicity>として、今夜あなたたちがワールドカップで優勝することを願っています。

Translate Test Data

[CODEC - Le et al., 2024]

今天是我的生日，却偏偏发现自己得了乳腺癌。

Today is my birthday, yet I've just discovered—of all days—that I have breast cancer.

<dob>Today is my birthday</dob>, yet I've just discovered—of all days—that <health>I have breast cancer</health>.

Today - How to Improve LLMs on Multilingual Applications?

Learning from Multilingual Human Preference
(*CARE* - EMNLP 2025)

Knowledge is not evenly distributed across languages.

offline reinforcement learning → **online reinforcement learning**

Multilingual Reinforcement Learning
(*LRPO* - ICML 2026)

Language/topic-aware routing and calibrated cross-lingual rewards are key to multilingual RL.

To Translate, or not to Translate — that's the Question

Translating training and/or test data offers a scalable and practical solution, with a few important caveats to consider. ⭐

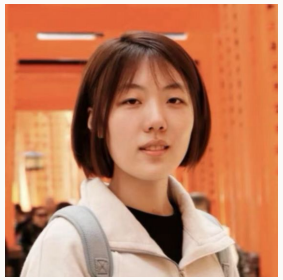
How do cross-lingual and cultural differences shape public opinions of AI?

Media Framing of AI

Julie Young

Katerina Addington

Yiren Wang



Geyang Guo

Tarek Naous

Anton Lavrouk

Will AI Save Us or Destroy Us? Comparing AI Media Framing Across Global Powers

Julie O Young **Katerina Addington** **Yiren Wang** **Geyang Guo** **Tarek Naous**

Anton Lavrouk **Zicong He** **Guanjun Yan** **Alan Ritter**

Wei Xu
Georgia Institute of Technology

Abstract

Public conversation about AI has accelerated since 2024, yet we lack cross-lingual, comparative evidence of how major news outlets across global powers frame AI. In this work, we study cross-national media framing of AI by constructing a multilingual dataset of 2,270 news articles published between January 2024 and October 2025 from 50 outlets across the U.S., China, and Russia. We introduce a taxonomy of 25 AI-specific frames with sentiment labels and perform sentence-level human annotation to capture how AI is discussed across languages. An analysis of our dataset reveals substantial cross-country differences: Chinese media consistently portray AI more positively, while U.S. and Russian coverage is more mixed; work-related frames dominate coverage, whereas safety-related concerns appear less frequently. We also observe frequent cross-national discourse, with media outlets often referencing other countries in competitive contexts. Using our data, we evaluate large language models (LLMs) on the framing classification task and show gaps in LLM performance, particularly in frame identification, highlighting the need for human-centered analysis of multilingual media narratives.

1 Introduction

As the use of AI skyrockets, public sentiment is tilting between anxiety and enthusiasm. A recent Pew Research Center [2025] survey across 25 countries found that people are twice as likely to feel concerned as excited about AI's growing role in daily life. While there are reasons to be optimistic, such as AI's potential to revolutionize scientific discovery [Bran et al., 2024, Park et al., 2025], there are also concerns from strain on the energy grid to automating away jobs [Bessen, 2018, Strubell et al., 2019]. These attitudes vary sharply by country. According to an Edelman [2024] poll, 87% of people in China say they trust AI technology, whereas only 32% of Americans do — an enormous gap! Meanwhile, as AI is rapidly integrated across virtually every industry, governments must grapple with appropriate regulations to mitigate AI-centric concerns while maintaining innovation in a hyper-competitive global environment. Against this backdrop, understanding how public perceptions of AI are shaped becomes increasingly critical.

Through this work, we seek to investigate *what the primary narratives are around AI as reflected in news media*, a key information source for the general public, across three major countries and their native languages: the United States, China, and Russia. While it is difficult to discern whether media influences audiences or simply mirrors public sentiment, identifying major topics of interest in media offers a grounded lens into how societies talk about AI and the national perspectives. What gets emphasized, as opportunity or threat, can shape how governments regulate AI, what technologies to fund and prioritize, and how citizens respond to AI's growing presence in their lives.

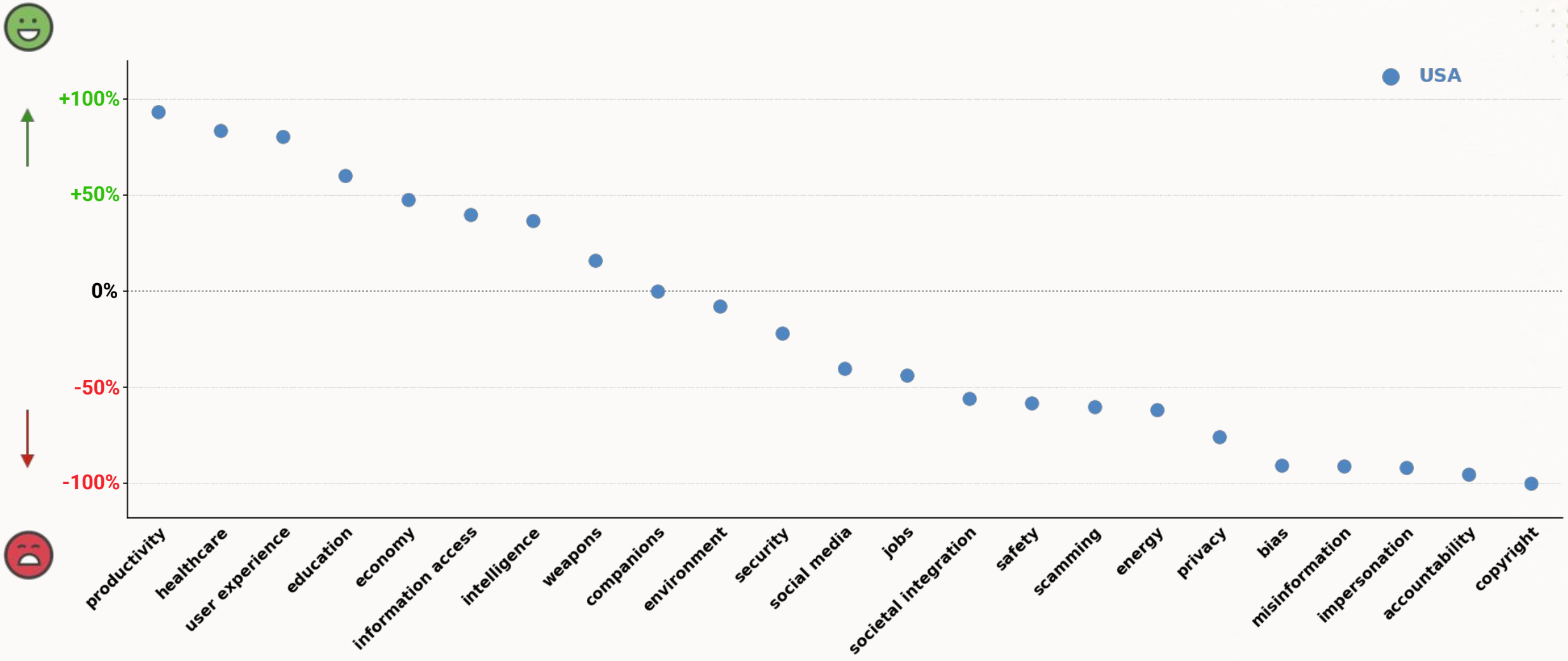
How do cross-lingual differences shape public perceptions?

We analyzed 2,270 articles from 50 media outlets in 3 countries.

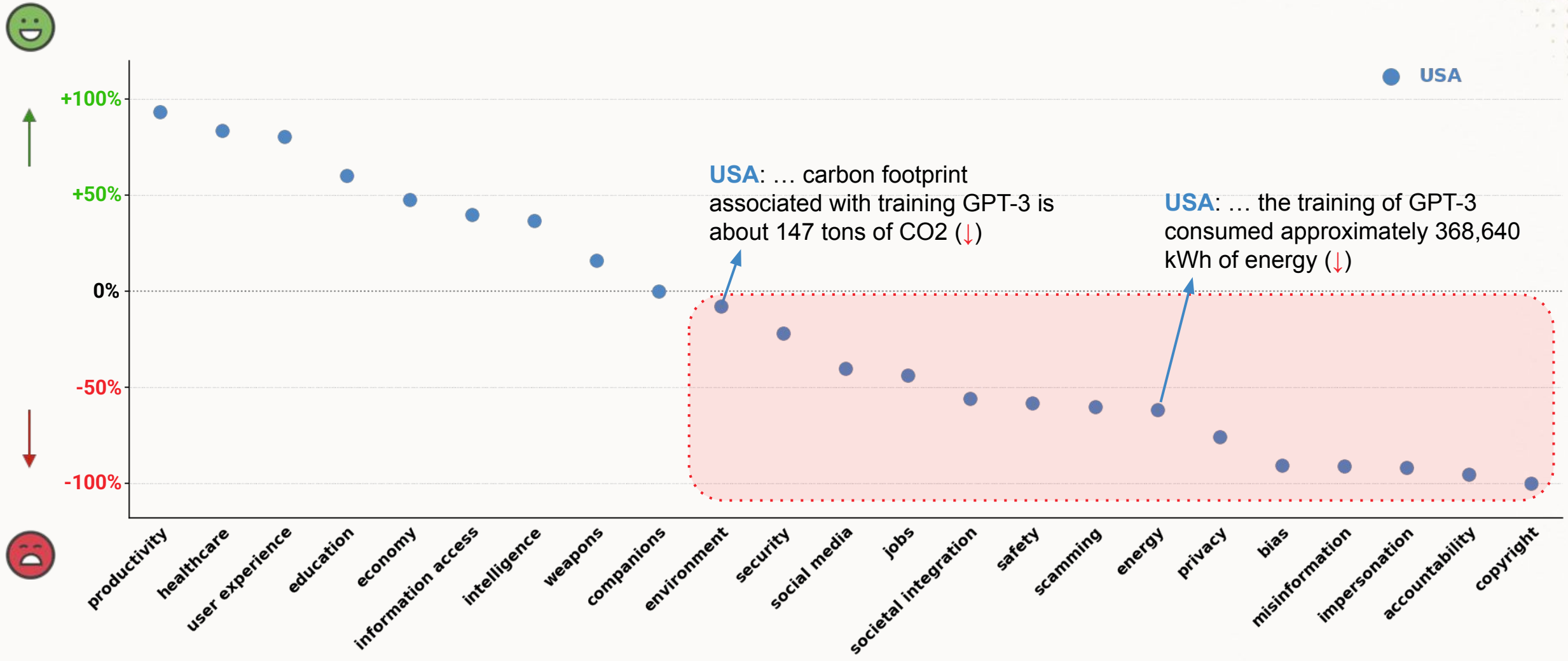
	USA	China	Russia
Traditional Media	New York Times Washington Post USA Today Forbes The Atlantic Politico The AP Newsweek Wall Street Journal The Hill Axios	People's Daily China News Service Xinhua Chengdu Business Daily The Beijing News Securities Times China Business Network China Times News Commercial Times NOWnews Va Kio Daily Ta Kung Pao	The Moscow Times Gazeta Russian Government RIA Novosti Lenta Economics and Life Metro News Pravda Severa Vechernaya Moskva Province VZ URA Federal Press The Insider EuroNews
Tech Media	The Verge MIT Tech Review VentureBeat TechRadar	TechNode TMTPost GuruClub	3DNews Cnews Hacker Mobile Review RB

How do cross-lingual differences shape public perceptions?

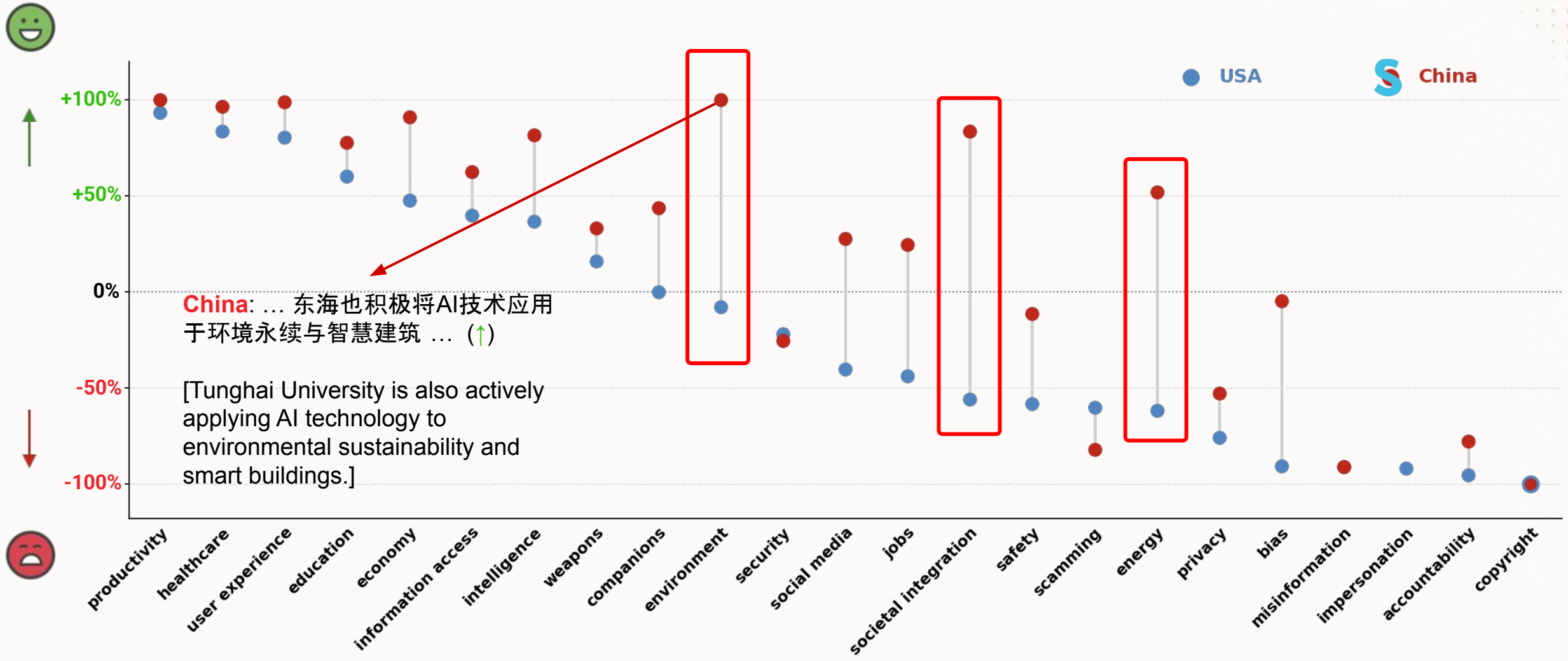
How do cross-lingual differences shape public opinions?



How do cross-lingual differences shape public opinions?



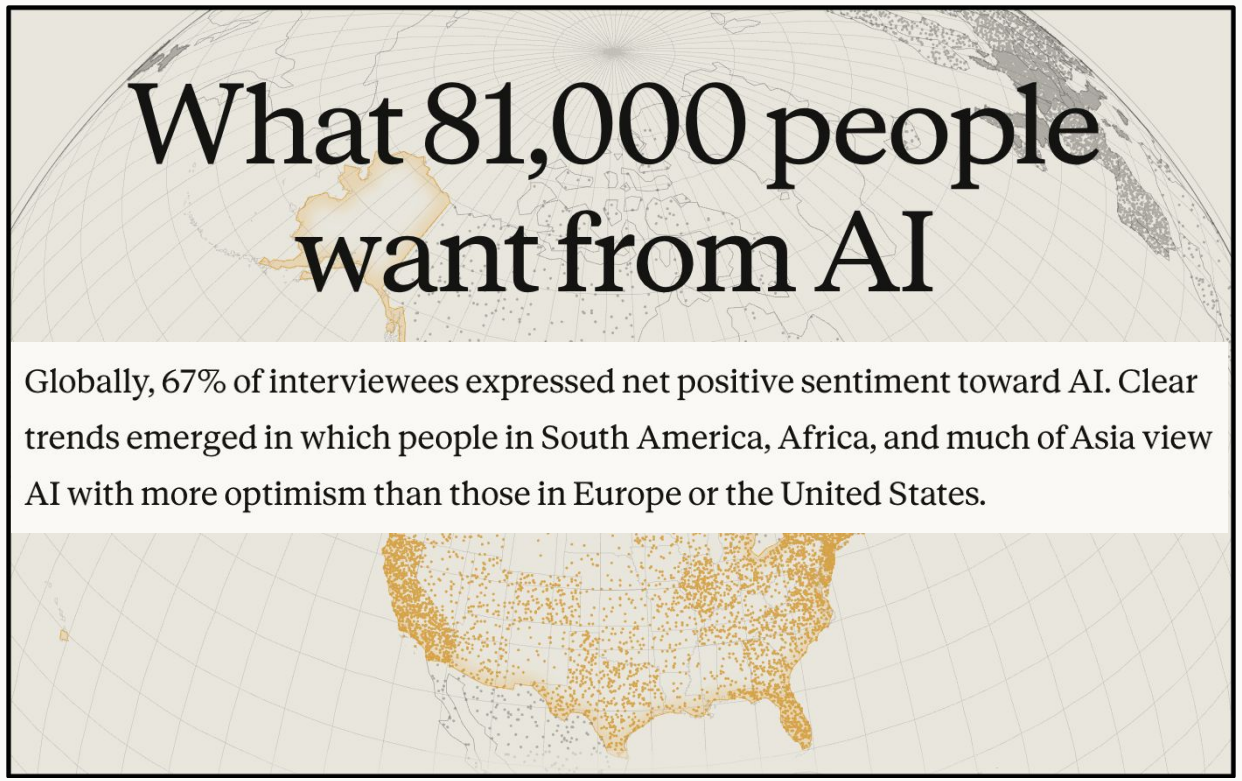
How do cross-lingual differences shape public opinions?



Mentions of other countries and regions in AI-related media coverage from the USA, Russia, and China.

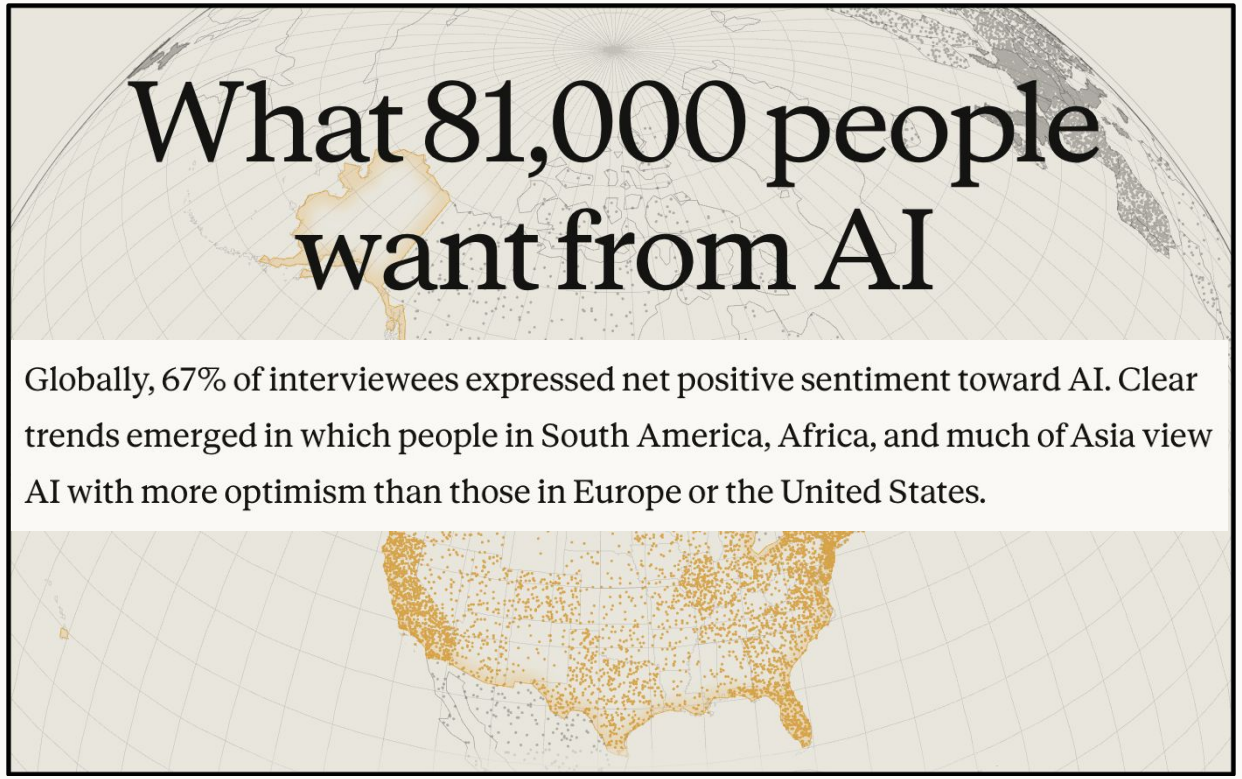
Category	Distribution	Top-2 Mentions	Examples
Competition	33%	China: 70%, Japan: 7%	"... the USA will lose ground in its AI race with **China**."
	29%	USA: 77%, ASEAN: 4%	"... while **the US** holds the leading position, mainland China is demonstrating astonishing speed in catching up."
	24%	USA: 41%, China: 39%	"... improved AI capabilities will strengthen its position against Apple and **Chinese** manufacturers."
Conflict	20%	China: 41%, Taiwan: 6%	"Meta says **the Kremlin** is the No. 1 source of AI-created misinformation."
	9%	USA: 100%	"... **the US** had already restricted companies exporting AI chips due to concerns that products might be diverted to China."
	13%	USA: 35%, China: 19%	"... sanctions imposed in **Ukraine** against Russia, **Belarus**, and **China**, ... production of drones using AI."
Collaboration	19%	China: 18%, UK: 15%	"**British** officials said the prime minister also hoped to announce cooperation ... including AI, with Mr. Trump."
	17%	USA: 31%, EU: 19%	"AI is... a crucial driving force for future China-**ASEAN** cooperation."
	39%	USA: 31%, China: 11%	"... joint efforts of Russian and **Chinese** scientists, it has been possible to create ... processors for AI tasks."
Innovations	28%	China: 21%, Japan: 10%	"Nissan has built the Yokohama Lab research center in **Japan** to study how to use AI in automotive manufacturing."
	46%	USA: 81%, EU: 7%	"Scientists used an AI system to analyze health data from hundreds of thousands of volunteers in the **UK** and the **US**."
	24%	USA: 40%, China: 14%	"**Ukraine**, **NATO countries**, **China**, and the U.S. are already implementing AI for autonomous drone control."

Russian media most frequently reference China and the USA, while Chinese and U.S. media primarily focus on each other.



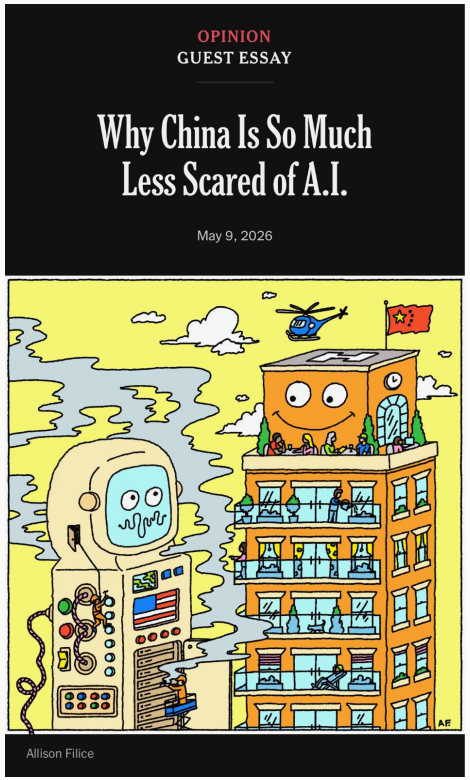
Perspectives on AI vary across regions

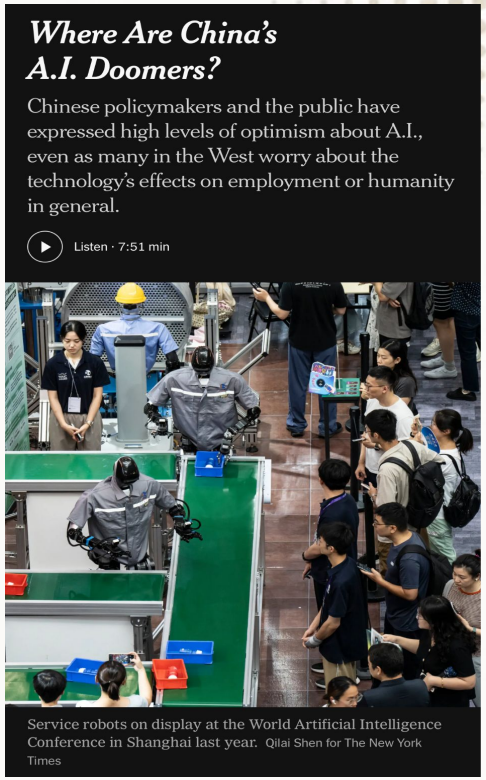
[Anthropic, 2026]



Perspectives on AI vary across regions

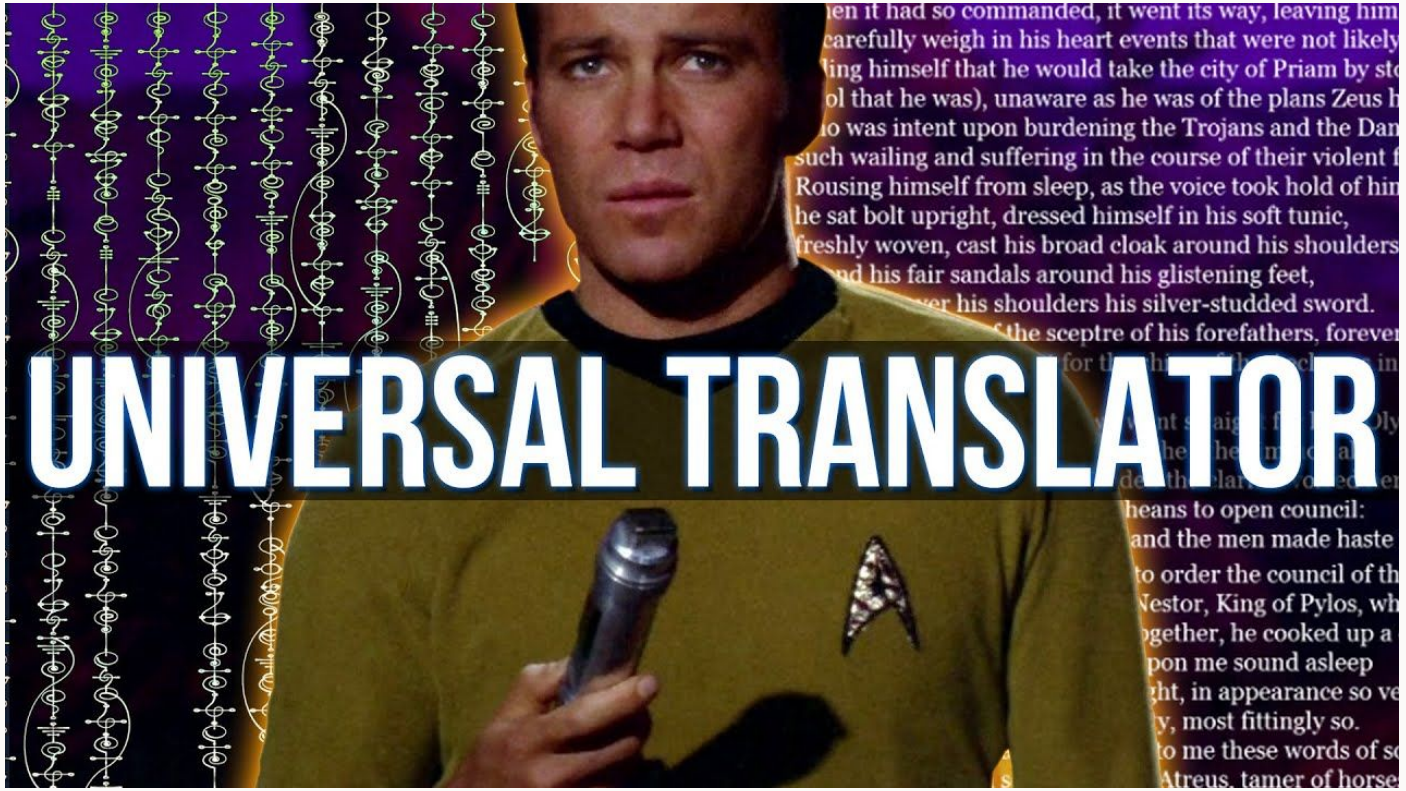
[Anthropic, 2026]





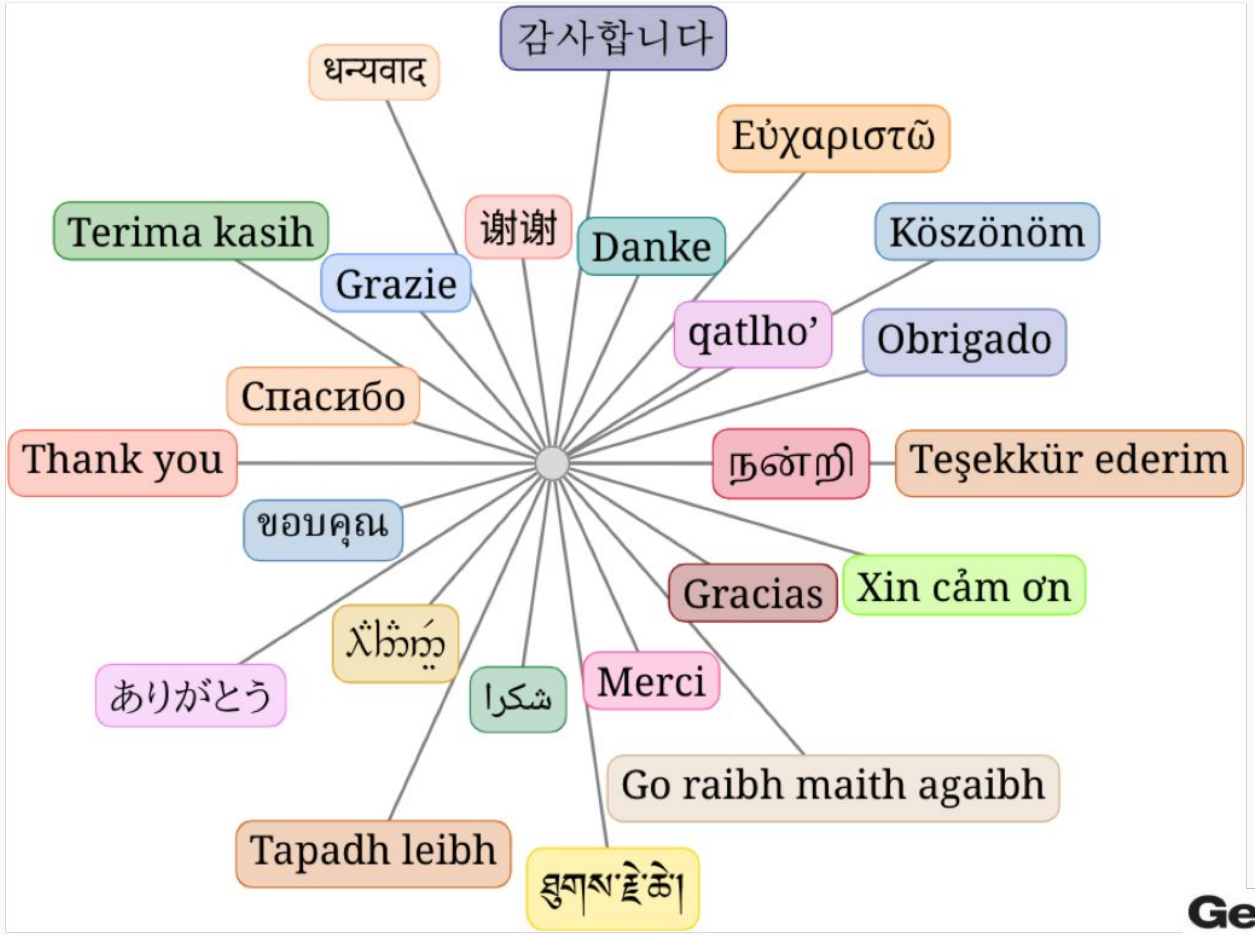
More and more reports on AI optimism in China, and Asia in general

Imaging a world with perfect translations

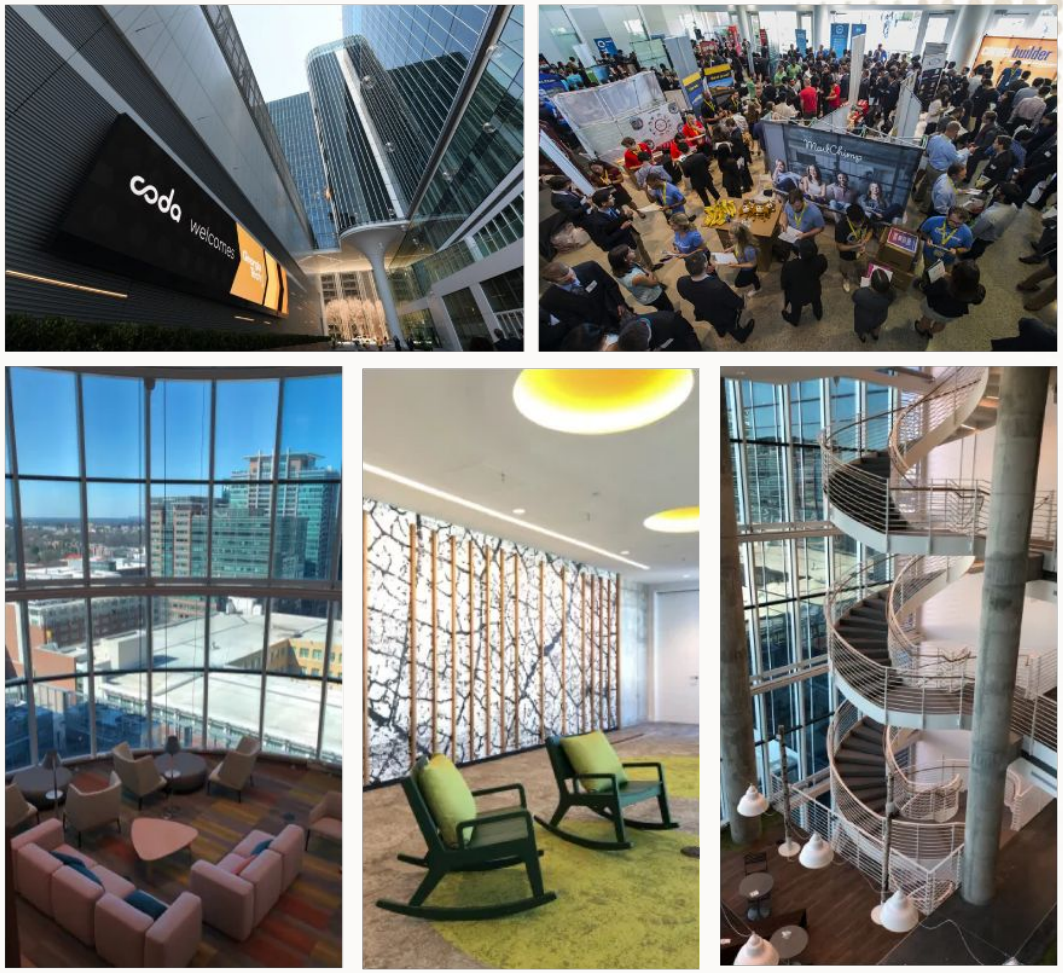


Thank you!

https://cocoxu.github.io/



(image credit: Overleafl)



(image credit: Georgia Tech)