Qingqiu-MT-9B: An Instruction-Following Multilingual Translation Model

Tian Xia **Chao Chen** **Mengpeng Yang** **Jingxu Yang** **Yabo Sun** **Qiang Liu**
Kingsoft Office WPS-Qingqiu / Wuhan, China
xiatian3@wps.cn, chenchao9@wps.cn, yangmengpeng@wps.cn
yangjingxu1@wps.cn, sunyabo@wps.cn, liuqiang2@wps.cn

Abstract

We present Qingqiu-MT-9B, an instruction-following multilingual translation model developed for the constrained track of the WMT26 General Machine Translation shared task. Our model supports translation among 20 languages and covers 22 of the 23 official language pairs. It uses a source-grounded synthesis pipeline with separate LLM-based steps for generating instance-specific translation instructions and their corresponding translations. The resulting corpus covers diverse translation requirements, task complexity, instruction formulations, domains, document lengths, and glossary constraints, reducing reliance on fixed prompt templates and supporting generalization to unseen translation instructions. Our model is based on Qwen3.5-9B and trained through two-stage full-parameter supervised fine-tuning (SFT) followed by reinforcement learning (RL). The first SFT stage performs broad adaptation on a large-scale, medium-quality corpus, while the second refines the model on a smaller, higher-quality set. The RL stage further optimizes the model with Group Relative Policy Optimization (GRPO). During development, we use an LLM-based, rubric-guided contrastive method that scores candidates by dimension and aggregates an overall score to assess translation quality. Experiments show that Qingqiu-MT-9B demonstrates strong generalization and competitiveness among similarly sized models. We release the model at https://huggingface.co/WPS-Qingqiu/Qingqiu-MT-9B.

1 Introduction

The WMT26 General Machine Translation Shared Task[1] covers 23 official language pairs. Unlike conventional translation settings that provide only a source text and target language, each sample includes a natural-language translation instruction. A system must therefore not only convey the source meaning accurately but also follow the translation requirements specified for each instance. This setting requires both robust multilingual translation capabilities and the ability to generalize to open-ended natural-language translation instructions.

For data construction, we do not treat instruction-following translation as a small-scale auxiliary task; instead, we make it a central component of our large-scale synthetic data. Accordingly, we construct a training corpus that pairs diverse source texts and translation instructions with high-quality target translations, thereby improving the model's ability to generalize across translation scenarios and to unseen translation instructions. Each synthesized example follows a shared schema containing the source language, target language, source text, a natural-language translation instruction grounded in the source, and its corresponding translation.

To broaden coverage of both source texts and translation instructions, we increase data diversity in source sampling and translation-instruction synthesis. Source texts are sampled by language, domain, and length. We then set the training-data proportions for each target-language direction and convert a randomly selected subset to HTML for end-to-end HTML translation. For translation-instruction synthesis, we factorize natural-language instructions along task complexity and instruction formulation, drawing their content from a carefully designed pool of translation requirements. Dynamic composition of the synthesis prompt controls how these factors are combined and distributed across the corpus. For each source text, an LLM is then used to generate a source-grounded and actionable natural-language translation instruction; once the instruction is fixed, a corresponding translation is generated according to its requirements. Compared with fixed-template synthesis, this factorized construction expands the combinations of translation requirements and instruction formula-

[1] https://www2.statmt.org/wmt26/translation-task.html

Proceedings of the Eleventh Conference on Machine Translation, pages 1447–1462
October 28-29, 2026 ©2026 Association for Computational Linguistics

tions, encouraging the model to learn translation behaviors that generalize across instruction formulations rather than surface patterns tied to specific prompt templates.

For model training, we build on Qwen3.5-9B (Qwen Team, 2026) with two-stage SFT followed by RL. The SFT stages are stratified by data quality. Stage 1 prioritizes scale and coverage, using large-scale, medium-quality multilingual data to establish foundational instruction-following translation capabilities. Stage 2 instead prioritizes quality. Data generated by a stronger teacher model are filtered to form a smaller, higher-quality training set, which is used to further improve translation quality. The RL stage then applies GRPO with contrastive LLM-based rewards (rubric-guided LLM-as-a-judge with language gates, plus GEMBA-ESA-style scoring) and rule-based length and format signals, using group-relative advantages over sampled outputs for the same prompt.

The main stages of our model development pipeline are:

- A source-grounded pipeline that combines multidimensional source sampling with factorized synthesis of translation instructions to create diversity across languages, domains, document lengths and formats, translation requirements, levels of task complexity and instruction formulations, and glossary constraints;
- A quality-tiered two-stage full-parameter SFT recipe that first learns from large-scale, medium-quality data and then refines the model on a smaller, higher-quality curated set;
- An RL stage using GRPO with contrastive LLM-based rewards (rubric-guided LLM-as-a-judge with language gates, plus GEMBA-ESA-style scoring) and rule-based length and format signals;
- An LLM-based, rubric-guided contrastive evaluation that scores candidates by dimension and aggregates an overall score to assess translation quality.

2 Related Work

Adapting large language models for translation. Work on LLM-based machine translation (MT) progresses from eliciting latent multilingual ability with prompts to translation-specific specialization. PaLM shows strong few-shot translation but still trails supervised MT systems (Vilar et al., 2023), while a 204-language study finds ChatGPT competitive mainly on higher-resource languages (Robinson et al., 2023). Translation-specialized open models then narrow this gap: ALMA combines monolingual continued pretraining with high-quality parallel refinement (Xu et al., 2024a), Tower adds multilingual instruction tuning (Alves et al., 2024), and ALMA-R introduces preference optimization (Xu et al., 2024b). WMT24 and WMT25 results place adapted LLMs among leading systems while showing that harder and lower-resource directions remain substantially challenging (Kocmi et al., 2024a, 2025). More recently, TranslateGemma specializes Gemma 3 through SFT on human and high-quality synthetic parallel data followed by RL with an ensemble of translation rewards, and reports consistent gains over its base models across broad multilingual evaluation (Finkelstein et al., 2026). The field thus shifts from scale alone toward data quality, multilingual coverage, and translation-specific post-training.

Supervised fine-tuning for MT. SFT is the standard first stage for adapting decoder-only LLMs to MT, with the source, language information, and translation requirements forming an instruction and the target translation providing token-level supervision. ParroT combines translation, contrastive, and error-guided instructions (Jiao et al., 2023), while multilingual translation-instruction tuning can transfer to unseen directions (Li et al., 2024). Recent work places greater emphasis on mixture design. LMT identifies directional degeneration when symmetric multiway data overemphasize many-to-one directions and mitigates it through strategic downsampling (Luo et al., 2025). MiLMMT-46 studies broad multilingual continued pretraining followed by a smaller, high-quality instruction-tuning set and finds strong many-to-many gains from moderate amounts of curated supervision (Shang et al., 2026). These results motivate attention to instruction diversity, direction balance, and target quality rather than raw bitext volume alone.

Reinforcement learning for MT. RL is used to reduce the mismatch between token-level maximum-likelihood training and sequence-level MT evaluation. Early approaches optimize metrics such as BLEU with REINFORCE or actor–

critic training (Ranzato et al., 2016; Bahdanau et al., 2017), while large-scale MT studies identify reward design and optimization stability as key concerns (Wu et al., 2018). More recently, GRPO becomes attractive for LLM-based MT because it derives relative advantages from multiple translations sampled for the same prompt without training a separate value model (Shao et al., 2024). Shy-hunyuan-MT applies GRPO after supervised distillation with a composite reward combining BLEU, XCOMET, and an LLM-based fluency score (Zheng et al., 2025). GRRM further argues that isolated scalar metrics may provide weak intra-group resolution and instead learns to score the full candidate group jointly (Yang et al., 2026). Conversely, WALAR shows that RL can amplify blind spots in source-based QE into omissions, over-translation, or wrong-language output (Liu et al., 2026). Reward design is therefore central to RL for MT and benefits from combining quality estimates with directly observable task constraints.

Automatic translation evaluation and quality estimation. Automatic MT assessment includes both reference-based metrics and reference-free quality estimation (QE). COMET learns from human judgments over the source, hypothesis, and reference (Rei et al., 2020), and CometKiwi extends this family to sentence- and word-level QE (Rei et al., 2022). MetricX-23 trains reference-based and QE regressors on human judgments and synthetic data (Juraska et al., 2023). MetricX-24 extends this family with a hybrid model and targeted synthetic failure cases (Juraska et al., 2024); MetricX-25 updates the family using Gemma 3 and introduces GemSpanEval for predicting error spans with categories and severities (Juraska et al., 2025). Prompted LLM judges form a complementary line. GEMBA scores translations with or without a reference (Kocmi and Federmann, 2023), while GEMBA-ESA follows the ESA protocol of combining error spans with holistic scoring (Kocmi et al., 2024b; Freitag et al., 2024). Candidate-aware evaluation is also emerging: COMET-poly contextualizes a hypothesis with alternative translations (Züfle et al., 2025), whereas Nvidia-Nemo jointly evaluates all candidates and finds relative evaluation effective when organized as a two-stage evaluate-then-refine procedure (Yan et al., 2025). WMT25 results nevertheless show that references still help at the segment level and that robustness across languages remains challenging (Lavie et al., 2025).

3 Task and System Overview

Figure 1 summarizes the system. Heterogeneous multilingual sources are first paired with source-grounded translation instructions, optionally including terminology constraints. A teacher then translates each source under its instruction. Failure-oriented filtering removes malformed, wrong-language, repetitive, or structurally inconsistent outputs. The resulting examples are converted to a two-message chat format. A large, broadly filtered corpus feeds Stage 1 SFT, while a smaller and more stringently curated high-quality set feeds Stage 2. The resulting policy is further optimized using the composite reward in Section 3.2.2.

3.1 Data Construction

Our data construction comprises three sequential stages: broad-coverage monolingual corpus collection, synthesis of diverse translation instructions, and high-quality translation synthesis. The three stages share one objective: at the instance level, each translation instruction should be grounded in its source and actionable, with its requirements realized by the target translation; at the corpus level, the resulting instruction-translation data should be large-scale, diverse, and high-quality.

3.1.1 Broad-coverage monolingual corpus collection

We draw source texts from heterogeneous multilingual pools. Public sources include MADLAD-400 (Kudugunta et al., 2023), FineWeb2 (Penedo et al., 2025), Common Crawl Traditional Chinese (Li, 2024), and MNBVC (MOP-LIWU Community and MNBVC Team, 2023). We complement these resources with the domain-labelled IndustryCorpus2 collection (Shi et al., 2024) and our in-house datasets. The combined pool spans everyday content as well as literary, news, technical, and financial domains. We retain longer passages in addition to sentence-level examples so that the model learns to follow translation instructions in document-like contexts.

Sampling is organized by language, domain, and text length. After sampling, we set the training-data proportions for each target-language direction and randomly convert a subset of texts to HTML to support end-to-end HTML translation.

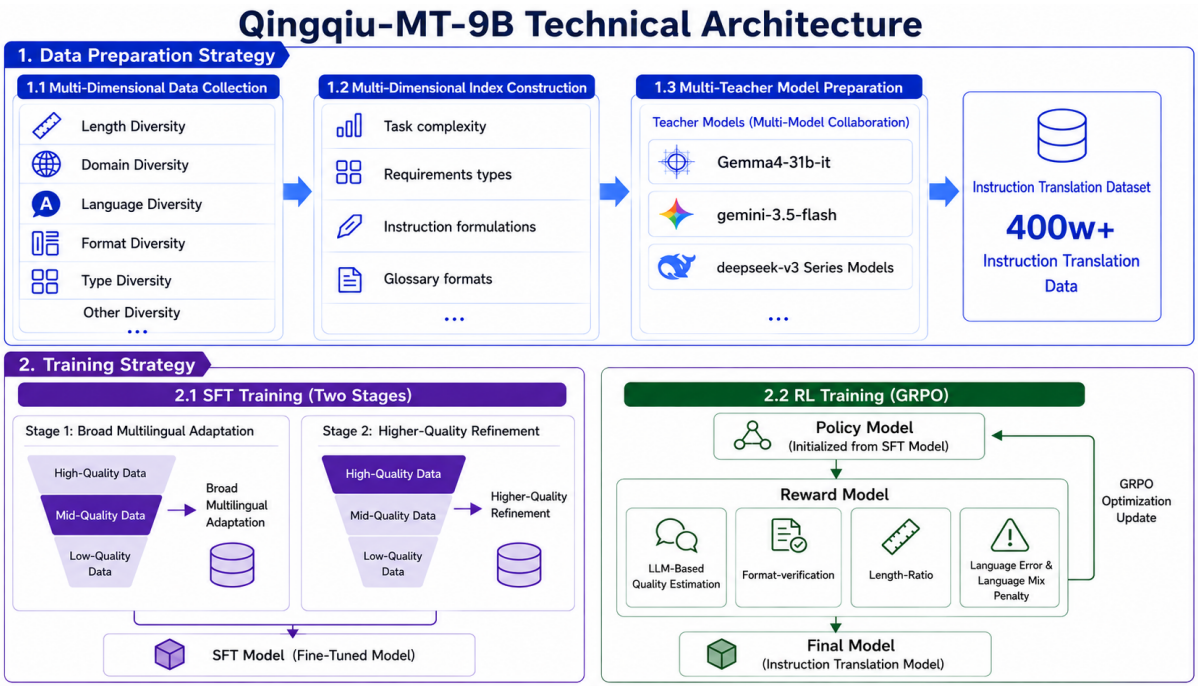

Figure 1: Overview of the Qingqiu-MT-9B architecture, including instruction-translation data construction, quality-tiered two-stage SFT, and GRPO-based training.

3.1.2 Synthesis of diverse translation instructions

For each source text, we synthesize a source-grounded and actionable natural-language translation instruction using the text and available metadata. Throughout this paper, an instruction denotes a translation task specification that governs how the source text should be translated; it excludes non-translation transformations such as rewriting, polishing, and expansion. The synthesis prompt provides a carefully designed, open-ended reference space of additional translation requirements, summarized in Table 1.

We factorize natural-language translation instructions along task complexity and instruction formulation, and further control glossary presentation when terminology constraints are present. We dynamically compose the synthesis prompt from selected additional requirements, a sampled complexity setting, a surface-style specification, and associated control information, and then use an LLM to generate a natural, coherent instruction that instantiates this configuration. This mechanism controls the corpus-level distribution of requirement types, levels of task complexity, instruction formulations, and glossary formats, increasing the coverage and diversity of the resulting instruction corpus.

For terminology-oriented translation tasks, source-language terms and their prescribed target-language equivalents are incorporated into the translation instruction using one of the glossary formats in Table 1. Terminology requirements can coexist with other translation requirements in the same instruction.

3.1.3 High-quality translation synthesis

Once the translation instruction is fixed, a teacher model takes the source language, target language, translation instruction, and source text as input and synthesizes the corresponding translation.

We filter synthetic outputs with missing fields, source copying, non-terminating repetition, wrong-language output or script mismatch, prompt leakage, or unsupported structural markup, followed by lightweight paragraph and format normalization.

We evaluated multiple combinations of synthesis prompt templates and teacher models, jointly considering synthesis quality, generation efficiency, and cost. Based on this comparison, we use Gemma4-31B and the DeepSeek-V3 series as teachers for Stage 1, and Gemini-3.5-Flash as the teacher for Stage 2.

Stage 1 contains approximately 4 million examples. Stage 2 uses approximately 500,000 examples selected under stricter criteria, emphasizing higher-quality translations and task-relevant translation instructions. The RL stage uses a separate curated set of approximately 50,000 examples paired with

Design dimension	Specification in the synthesis prompt
Additional translation requirements	*Open-ended candidate pool.* register/formality; tone/emotion; translation strategy; domain adaptation; purpose/scenario; punctuation styling; implicitation/implicitness preservation; additional context/background; fidelity–naturalness trade-off; audience adaptation; verbosity preference; localization/cultural rendering; and glossary/terminology constraints.
Task complexity	*Mixture.* single additional requirement (60%); multiple compatible requirements (40%; 2–5 items).
Instruction formulations	*Predefined set.* direct_imperative; polite_request; scenario_based; audience_first; editor_brief; localization_memo; concise_checklist; translator_note; constraint_focused; and comparative.
Glossary formats	*Predefined set.* arrow (`s` → `t`); parenthetical (`s (t)`); translate_as (`Translate "s" as "t"`); use_for (`Use "t" for "s"`); dict_style (`{"s":"t"}`); list_of_pairs (`[("s","t")]`); pipe (`s | t`); slash (`s / t`); and equals (`s = t`).

Table 1: Prompt-controlled design space for translation-instruction synthesis. Additional translation requirements are drawn from an open-ended candidate pool; task complexity, instruction formulations, and glossary formats are controlled by a predefined mixture or style set.

translations.

3.2 Model Training

Our training procedure follows a three-stage curriculum comprising two quality-tiered SFT stages followed by RL. Stage 1 favors scale and coverage with a large, moderately filtered corpus, whereas Stage 2 favors supervision quality with a much smaller and more stringently curated set. This ordering first establishes broad translation competence and then refines it using a denser, cleaner training signal.

3.2.1 Two-stage supervised fine-tuning

We initialize from Qwen3.5-9B (Qwen Team, 2026). Because the task is text-only, we update all language-model parameters while freezing the visual components. Both SFT stages use a chat template, with prompts composed in the same way as during final inference.

Stage 1: broad multilingual adaptation. The first stage uses the large corpus described in Section 3.1 to establish broad translation-direction coverage, robustness across domains and text lengths, and foundational instruction-following translation capabilities. In addition to the 22 translation directions specified by the WMT task, we augment the training data with reverse directions and other directions among the system's supported languages, and set the mixture proportions by target language. This design promotes more general linguistic knowledge and cross-lingual translation behavior. Careful mixture design also reduces interference among translation tasks and prevents data-rich directions from crowding out others during training. Figure 2 presents the resulting mixture from two complementary views: translation-direction groups and target-language coverage.

Stage 2: higher-quality refinement. The Stage 2 dataset is restricted to the translation directions specified by the WMT task and is selected using stricter quality criteria. Although smaller than Stage 1, it contains higher-quality teacher translations and emphasizes examples involving complex translation instructions, terminology constraints, and style-sensitive translation. Building on the broad adaptation acquired in Stage 1, this higher-quality supervision can further improve the model's translation quality and ability to follow instructions on the specified WMT translation directions.

3.2.2 Reinforcement Learning with GRPO

After two-stage SFT, we apply a reinforcement-learning stage using GRPO (Shao et al., 2024) to further improve output quality and stability. Following GRPO, for each prompt q, the old policy $\pi_{\theta_{\text{old}}}$ samples a group of G outputs, $\mathcal{Y}_q = \{y_i\}_{i=1}^{G}$ (in our practice, a rollout count of $G = 4$ provides the best trade-off between training efficiency and performance). Each sampled output receives a composite reward organized into LLM-based and rule-based families: the two families estimate translation quality from different perspectives and induce discriminative relative score gaps among the sampled outputs. These gaps are normalized within the group to form the group-relative advantages used by GRPO, without training a separate value model. The LLM-based family contains a rubric-

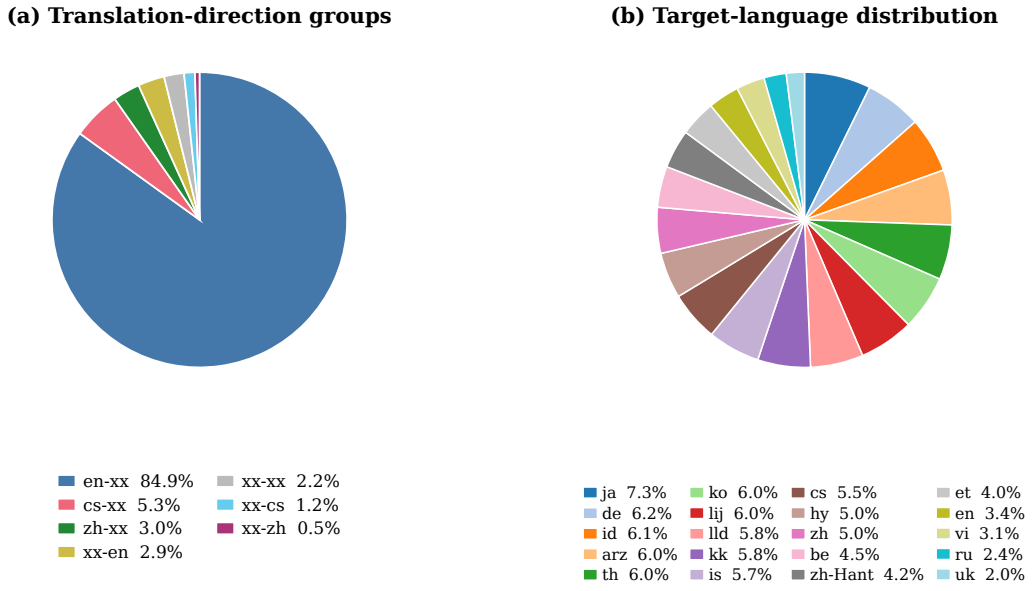

Figure 2: Distribution of training examples in the Stage 1 SFT corpus. (a) Mutually exclusive direction groups centered on English, Czech, and Chinese. Here, `xx` is a wildcard over supported languages; source-side `en`, `cs`, and `zh` groups take precedence, and the remaining examples are grouped by core target language. (b) Aggregation by target language across the 20 languages supported by the system. Percentages are computed from example counts.

guided LLM-as-a-judge score and an error-oriented GEMBA-ESA-style score; the rule-based family contains length-ratio and format-verification metrics. The LLM-as-a-judge score is further controlled by two gates for target-language correctness and language-mixing detection. We use a DAPO-style clipped objective (Yu et al., 2025) together with a small KL regularizer to constrain drift from the supervised policy.

The final scalar reward is expressed as the sum of the two reward families:

$$\begin{aligned} R(y) &= R_{\text{LLM}}(y) + R_{\text{rule}}(y), \\ R_{\text{LLM}}(y) &= 1.0 r_{\text{LLM}}(y) + 0.5 r_{\text{ESA}}(y), \\ R_{\text{rule}}(y) &= 0.5 r_{\text{len}}(y) + 0.5 r_{\text{fmt}}(y). \end{aligned} \tag{1}$$

All four component signals are normalized to $[0, 1]$.

Contrastive LLM-based rewards. The LLM-based family comprises two complementary contrastive rewards.

Rubric-guided LLM-as-a-judge scoring with dual language gating. We call GLM-5 (GLM-5 Team, 2026) to jointly evaluate all sampled outputs in $\mathcal{Y}_q$ in a single request. The sampled outputs are presented under anonymized labels without generation provenance, yielding a contrastive blind evaluation. Under a shared rubric, the judge compares translation accuracy, terminology, compliance with the translation instruction, fluency, style, and required structure, and assigns sampled output y_i a score $s_i \in [0, 100]$. We normalize this score and apply two binary gates to obtain

$$r_{\text{LLM}}(y_i) = g_{\text{tgt}}(y_i)\, g_{\text{mix}}(y_i) \frac{s_i}{100}. \tag{2}$$

The target-language gate g_{tgt} passes only when the output is consistent with the requested target language, while the mixing gate g_{mix} passes only when no unexpected language or script mixing is detected. Failing either gate sets r_{LLM} to zero.

This groupwise comparison directly matches GRPO's group-relative advantage estimation over sampled outputs for the same prompt. One groupwise request replaces G independent scoring requests, reducing the number of translation-quality judge calls to $1/G$ of per-output evaluation. Placing the entire output group in the same judging context also improves the stability of within-group scores.

Error-oriented GEMBA-ESA-style scoring. Because the original GEMBA-ESA protocol (Kocmi et al., 2024b; Freitag et al., 2024) does not condition on translation instructions, we adapt it to an instruction-aware variant for our setting. Like the rubric-guided judge, this GEMBA-ESA-style evaluator jointly scores all anonymized sampled outputs in the group in a single call: along the dimensions of accuracy, fluency, style, terminology,

and non-translation, it compares the sampled outputs and, for each one, reports error types with their severities together with a 0–100 score. We normalize that score to $r_{\mathrm{ESA}} \in [0, 1]$. The rubric-guided LLM-as-a-judge reward provides positive scoring of relative translation quality, whereas the GEMBA-ESA-style reward lowers scores from an error-oriented perspective. The two contrastive LLM-based rewards therefore provide complementary positive-scoring and negative-deduction signals.

Rule-based rewards. The rule-based family comprises a length-ratio metric and a format-verification metric.

Length-ratio metric. In practice, we find that the length-ratio metric effectively discourages severe omissions and additions, as well as unbounded generation. Reflecting the nature of translation, we allow output lengths to vary within a tolerance band rather than forcing them to match the reference. Let $\rho = N(y)/N(y^*)$ be the output-to-reference token ratio and $\delta = \max(0, |\log \rho| - |\log 0.7|)$. We use

$$r_{\mathrm{len}}(y) = \frac{1}{1 + 30\delta^2}, \tag{3}$$

which has a broad full-score region and decays smoothly and rapidly when the length discrepancy becomes large.

Format-verification metric. We first classify the source with rule-based heuristics and route it to a type-specific check: for JSON, we verify that the sampled output parses correctly and matches the source's element count; for HTML-tagged content, we check whether the sampled output and the reference agree in newline and HTML-tag counts; for plain text, we check agreement in newline counts. The metric is binary: $r_{\mathrm{fmt}} = 1$ if the check passes and 0 otherwise.

3.3 Implementation

All models are implemented in PyTorch and trained on 16 GPUs. We train all stages with the ms-swift framework (Zhao et al., 2025). Table 2 summarizes the main hyperparameters for SFT and GRPO.

Hyperparameters	SFT_1	SFT_2	GRPO
Learning rate	2e−5	2e−6	2e−6
Batch size	384	64	128
Max sequence length	4096	4096	2048
Number of generations	–	–	4
Temperature	–	–	1.0

Table 2: Main training hyperparameters for SFT and GRPO.

4 Evaluation

4.1 Evaluation Datasets

We report development results on IF-WMT25 and IFMTBench (Sun et al., 2026). IF-WMT25 is our primary development set and covers 22 of the WMT26 translation directions (excluding English→Northern Sámi): Czech and English sources are taken from the WMT25 blind test (Kocmi et al., 2025), Chinese sources are sampled from the same source pool described in Section 3.1, and translation instructions are generated with the instruction-synthesis procedure in that section. The resulting set contains 5,239 examples, and its instructions are English-only. IFMTBench is an open benchmark for instruction-following MT covering seven languages, six constraint types, and five compositional patterns. We evaluate on its multi-constraint subset of 2,838 examples, whose translation instructions are multilingual and whose language pairs are not restricted to the WMT26 inventory; we use it to assess generalization to non-English instructions and beyond the official language-pair inventory.

4.2 Evaluation Setup

Baselines. We compare Qingqiu-MT-9B separately with Qwen3.5-9B (Qwen Team, 2026), Hy-MT2-7B (Zheng et al., 2026), Gemma4-12b (Gemma Team and Google DeepMind, 2026), and DeepSeek-v4-flash (DeepSeek-AI, 2026).

Inference settings. For Qingqiu-MT-9B, we reuse the training chat template, map language codes to names, and decode in non-thinking mode to return translation text only. Paragraphs are serialized and restored with `\n\n` under one fixed decoding configuration for all directions. Although trained with English instructions only, we keep IFMTBench's original multilingual instructions at evaluation. For the baselines, we decode in non-thinking mode with the recommended sampling settings from their official documentation. On IFMTBench, they use the input prompts provided by the benchmark; on IF-WMT25, they use the same input prompts as Qingqiu-MT-9B.

Evaluation method. For each source and instruction, GLM-5 jointly scores anonymized system outputs on five $[0, 100]$ dimensions—accuracy, terminology, format, fluency, and style—folding instruction following into these dimensions rather than scoring it separately. The overall score is the rounded weighted sum of the five scores with weights $(0.40, 0.20, 0.15, 0.15, 0.10)$; absolute differences of at most 2 count as ties. We use this method for both evaluation sets in place of the original IFMTBench scoring. Tables report win and lose rates (%) against each baseline, together with $\text{Gap} = \text{Win} - \text{Lose}$; ties stay in the denominator but are omitted from the display. Positive gaps are shown in red and negative gaps in green. Direction rows are computed over samples in that direction.

Because GLM-5 also serves as an RL judge, reusing it for development evaluation raises the possibility of judge-specific reward hacking. To test the robustness of our conclusions to judge choice, we conduct a judge ablation on IF-WMT25. DeepSeek-V4-Flash-Thinking (High) and GPT-5.6-Sol-Thinking (High) are separately used as judges under the same instruction-aware evaluation rubric. Table 5 compares their results with the original GLM-5 evaluation.

4.3 Results

Tables 3 and 4 report pairwise win rates. We treat IF-WMT25 as the primary development evaluation because it covers the WMT26 translation directions.

	Qingqiu-MT-9B / baseline (%)											
	Qwen3.5-9B			Hy-MT2-7B			Gemma4-12b			DeepSeek-v4-flash		
Direction	Win	Lose	Gap	Win	Lose	Gap	Win	Lose	Gap	Win	Lose	Gap
cs→de	81.1	18.9	62.2	45.2	52.4	-7.2	44.8	48.0	-3.2	29.9	63.8	-33.9
cs→uk	84.1	15.1	69	43.9	52.0	-8.1	53.4	42.5	10.9	29.9	63.8	-33.9
cs→vi	80.3	17.3	63	61.6	36.8	24.8	46.6	40.6	6	46.5	48.0	-1.5
en→arz	93.7	6.3	87.4	78.7	18.1	60.6	35.6	54.9	-19.3	44.9	52.8	-7.9
en→be	86.7	12.5	74.2	–	–	–	82.8	12.8	70	29.5	62.3	-32.8
en→cs	90.5	7.9	82.6	39.1	57.0	-17.9	63.5	29.4	34.1	29.4	65.1	-35.7
en→de	86.1	12.0	74.1	56.1	43.0	13.1	34.6	55.7	-21.1	41.7	51.9	-10.2
en→et	88.0	10.4	77.6	–	–	–	81.1	16.9	64.2	27.2	68.8	-41.6
en→hy	96.1	3.9	92.2	–	–	–	76.4	22.4	54	41.1	57.3	-16.2
en→id	80.8	15.4	65.4	44.4	55.6	-11.2	37.5	57.8	-20.3	48.8	43.9	4.9
en→is	94.5	5.5	89	–	–	–	84.1	13.5	70.6	34.9	61.9	-27
en→ja	83.5	13.4	70.1	39.3	55.7	-16.4	31.5	52.2	-20.7	29.9	55.9	-26
en→kk	89.6	10.4	79.2	83.5	15.7	67.8	58.7	37.7	21	43.2	52.0	-8.8
en→ko	86.6	11.0	75.6	44.5	48.4	-3.9	31.0	61.9	-30.9	43.2	48.0	-4.8
en→lij	78.2	20.2	58	–	–	–	67.7	26.3	41.4	52.0	44.0	8
en→lld	70.2	27.4	42.8	–	–	–	79.7	17.6	62.1	53.7	39.7	14
en→ru	72.4	13.0	59.4	51.2	32.8	18.4	33.7	34.1	-0.4	35.7	40.5	-4.8
en→th	94.9	5.1	89.8	58.0	35.8	22.2	34.6	54.6	-20	66.7	26.9	39.8
en→uk	87.2	6.4	80.8	46.8	46.0	0.8	38.6	52.6	-14	29.5	57.4	-27.9
en→zh	62.7	26.2	36.5	46.0	37.1	8.9	41.0	41.4	-0.4	30.4	47.2	-16.8
en→zh-Hant	61.0	30.9	30.1	48.0	35.4	12.6	37.5	47.8	-10.3	33.1	48.8	-15.7
zh→ja	85.9	9.4	76.5	50.8	35.9	14.9	42.5	37.2	5.3	45.7	40.2	5.5
Overall	83.2	13.7	69.5	52.4	40.8	11.6	52.3	38.3	14	38.7	52.4	-13.7

Table 3: IF-WMT25 results against each baseline (%).

	Qingqiu-MT-9B / baseline (%)											
	Qwen3.5-9B			Hy-MT2-7B			Gemma4-12b			DeepSeek-v4-flash		
Direction	Win	Lose	Gap	Win	Lose	Gap	Win	Lose	Gap	Win	Lose	Gap
de→en	75.0	0.0	75	37.5	37.5	0	12.5	37.5	-25	25.0	25.0	0
en→de	63.6	27.3	36.3	54.5	45.5	9	18.2	72.7	-54.5	18.2	63.6	-45.4
en→ja	64.3	21.4	42.9	40.0	33.3	6.7	53.3	26.7	26.6	28.6	35.7	-7.1
en→ko	88.9	11.1	77.8	33.3	55.6	-22.3	55.6	22.2	33.4	22.2	77.8	-55.6
en→zh	43.5	34.8	8.7	42.6	31.3	11.3	30.8	29.1	1.7	23.1	41.0	-17.9
ja→en	50.0	8.3	41.7	16.7	33.3	-16.6	33.3	25.0	8.3	8.3	33.3	-25
ja→zh	65.6	6.2	59.4	56.2	14.6	41.6	50.0	14.6	35.4	36.5	27.1	9.4
ko→en	64.3	14.3	50	21.4	35.7	-14.3	21.4	35.7	-14.3	42.9	21.4	21.5
ko→zh	47.4	36.8	10.6	42.1	26.3	15.8	36.8	31.6	5.2	26.3	36.8	-10.5
zh→en	43.3	29.4	13.9	56.9	18.1	38.8	28.0	25.1	2.9	31.6	36.6	-5
zh→ja	84.6	15.4	69.2	46.2	30.8	15.4	30.8	46.2	-15.4	38.5	30.8	7.7
zh→ko	83.3	16.7	66.6	50.0	38.9	11.1	38.9	55.6	-16.7	44.4	44.4	0
Overall	52.5	26.6	25.9	45.9	28.5	17.4	31.9	30.8	1.1	28.3	38.8	-10.5

Table 4: IFMTBench multi-constraint results against each baseline (%).

On IF-WMT25 (Table 3), Qingqiu-MT-9B outperforms Qwen3.5-9B on all 22 directions, Hy-MT2-7B on 10 of 16 directions with available outputs, and Gemma4-12b on 11 of 22 directions,

	GLM-5				DeepSeek-V4-Flash-Thinking (High)				GPT-5.6-Sol-Thinking (High)			
Baseline	Win	Lose	Tie	Gap	Win	Lose	Tie	Gap	Win	Lose	Tie	Gap
Qwen3.5-9B	83.2	13.7	3.1	69.5	79.4	11.6	9.0	67.8	83.1	13.5	3.4	69.6
Hy-MT2-7B	52.4	40.8	6.8	11.6	40.7	36.9	22.4	3.8	51.0	40.4	8.6	10.7
Gemma4-12b	52.3	38.3	9.4	14.0	45.4	31.2	23.4	14.1	50.2	40.2	9.5	10.0
DeepSeek-v4-flash	38.7	52.4	8.9	-13.7	28.9	46.6	24.4	-17.7	37.5	52.1	10.3	-14.6

Table 5: Judge ablation on IF-WMT25. Win, Lose, Tie, and Gap are percentages for Qingqiu-MT-9B against each baseline; ties remain in the denominator, and $\mathrm{Gap} = \mathrm{Win} - \mathrm{Lose}$. GLM-5 values reproduce the Overall row of Table 3.

with higher overall win rates in each comparison. DeepSeek-v4-flash retains a higher overall win rate, although Qingqiu-MT-9B wins 5 of 22 directions, including lower-resource pairs such as en→th, en→lij, and en→lld.

The judge ablation exhibits the same overall trend across all three evaluators. Qingqiu-MT-9B consistently shows a clear advantage over Qwen3.5-9B and remains ahead of the similarly sized Hy-MT2-7B and Gemma4-12b, although the margins and tie propensities vary by judge. DeepSeek-v4-flash remains the strongest comparator, while Qingqiu-MT-9B retains a competitive position. The consistent ordering and direction of these comparisons indicate that the main conclusions are robust to evaluator choice and are unlikely to be explained by reward hacking against the GLM-5 judge.

On IFMTBench's multilingual instructions and out-of-inventory language pairs (Table 4), Qingqiu-MT-9B wins all directions against Qwen3.5-9B and achieves higher overall win rates than Hy-MT2-7B (8 of 12 directions) and Gemma4-12b (7 of 12). It wins 3 of 12 directions and ties 2 against DeepSeek-v4-flash, while trailing in overall win rate. Together, the two evaluations indicate generalization across instruction languages and translation directions.

4.4 Qualitative Analysis

Figure 3 shows four contrastive pairs aligned with the instruction phenomena highlighted on the WMT26 General MT task page (formal/informal, glossaries, structured translation, and style and expressions; two cases per type). Each prompt lists the language pair, instruction, and source text; colored spans mark the key constraint. Within a type, the source is held fixed so that differences in the output can be attributed to the instruction. Additional cases aligned with the design space in Table 1 appear in Appendix A.

formal / informal. For the campus dining-hall notice, a formal official-notice instruction yields impersonal bulletin Chinese with official-notice obligation markers, whereas an informal group-chat instruction produces colloquial wording for the same closing hours and self-brought bottle rule.

glossaries. The ambiguous sentence *Curry is best when it's hot* is resolved by additional context: as a food dish, the model translates *Curry*/*hot* as heated curry; as NBA player Stephen Curry, it renders a "hot streak" reading instead of food temperature.

structured translation. On the same password-reset HTML snippet, one instruction keeps tags and attributes such as `class` and `href` while translating only visible text; the contrastive instruction strips markup and returns plain Chinese only.

style and expressions. For an informal product comment with slang, character elongation, and exaggerated punctuation, one instruction replicates the non-standard emotional style while localizing slang into Chinese (e.g., rendering *tbh* rather than leaving the English token), whereas the contrastive instruction yields a polished product-review register.

5 Conclusion

We presented Qingqiu-MT-9B, an instruction-aware multilingual translation model trained with diverse synthetic supervision, quality-tiered two-stage SFT, and GRPO combining contrastive LLM-based and rule-based rewards. Its shared training and inference interface targets generalization beyond fixed prompt templates. Development results are encouraging on IF-WMT25, which covers the WMT26 directions, and remain competitive on IFMTBench under multilingual instructions and non-official language pairs.

Type	Prompt	Output
formal / informal	*[English → Simplified Chinese]* **Instruction:** Use a formal official-notice voice (公文通告体): impersonal and formulaic. Use obligation markers such as 须/请 where natural. Do not add dates or information absent from the source. **Source:** Campus dining halls will close at 8 p.m. during the exam week. Students should bring their own water bottles; disposable cups will not be provided.	考试周期间，校园食堂将于晚上8时关闭。学生须自带水杯，食堂不提供一次性纸杯。
	[English → Simplified Chinese] **Instruction:** Use a natural informal register, as if a student is casually telling classmates in a group chat. Keep it colloquial; avoid official-notice wording. **Source:** Campus dining halls will close at 8 p.m. during the exam week. Students should bring their own water bottles; disposable cups will not be provided.	考试周食堂晚上8点就关门了。大家记得自带水杯，不提供一次性杯子哈。
glossaries	*[English → Simplified Chinese]* **Instruction:** Additional context: this sentence is about the food dish curry; "hot" means served at a high temperature. **Source:** Curry is best when it's hot.	咖喱趁热吃最好。
	[English → Simplified Chinese] **Instruction:** Additional context: this sentence is about NBA player Stephen Curry; "hot" means he is on a scoring streak / shooting very well, not food temperature. **Source:** Curry is best when it's hot.	库里在手感火热时状态最好。
structured translation	*[English → Simplified Chinese]* **Instruction:** Keep all HTML tags and attributes intact (including class and href). Translate only visible text. **Source:** `<p class="alert">Please <a href="/reset">reset your password</a> before Friday.</p>`	请在周五之前 重置您的密码。
	[English → Simplified Chinese] **Instruction:** Remove all HTML tags and attributes. Output plain text only, with no '<' or '>' characters. **Source:** `<p class="alert">Please <a href="/reset">reset your password</a> before Friday.</p>`	请在周五之前重置您的密码。
style and expressions	*[English → Simplified Chinese]* **Instruction:** Accurately replicate the non-standard emotional style: keep character elongation (e.g. sooooo) and expressive punctuation (!!! / ???). Translate internet slang into natural Chinese equivalents; do not leave English slang tokens. **Source:** tbh this phone is sooooo good lol battery lasts forever!!! charging??? instant	说实话这手机真的超超超好用哈哈 电池简直能用到天荒地老！！！充电？？？秒充
	[English → Simplified Chinese] **Instruction:** Use a standard, polished product-review register. Normalize slang and elongation; conventional punctuation only. **Source:** tbh this phone is sooooo good lol battery lasts forever!!! charging??? instant	说实话，这款手机非常出色，电池续航极其强劲，充电速度也非常快。

Figure 3: Four contrastive instruction-following pairs (two cases per type): formal/informal, glossaries, structured translation, and style and expressions. Within each type, the source is held fixed while the instruction changes. Colored text highlights the key constraint.

Acknowledgments

We thank the Engineering and Infrastructure teams at WPS-Qingqiu for their support throughout this project. Their assistance with training infrastructure, large-scale experiment orchestration, and system reliability made the model development and evaluation reported in this paper possible.

References

Duarte Miguel Alves, José Pombal, Nuno M. Guerreiro, Pedro Henrique Martins, João Alves, M. Amin Farajian, Ben Peters, Ricardo Rei, Patrick Fernandes, Sweta Agrawal, Pierre Colombo, José G. C. de Souza, and André F. T. Martins. 2024. Tower: An open multilingual large language model for translation-related tasks. In *First Conference on Language Modeling*.

Dzmitry Bahdanau, Philemon Brakel, Kelvin Xu, Anirudh Goyal, Ryan Lowe, Joelle Pineau, Aaron C. Courville, and Yoshua Bengio. 2017. An actor-critic algorithm for sequence prediction. In *International Conference on Learning Representations*.

DeepSeek-AI. 2026. DeepSeek-V4: Towards highly efficient million-token context intelligence. *Preprint*, arXiv:2606.19348.

Mara Finkelstein, Isaac Caswell, Tobias Domhan, Jan-Thorsten Peter, Juraj Juraska, Parker Riley, Daniel Deutsch, Geza Kovacs, Cole Dilanni, Colin Cherry, Eleftheria Briakou, Elizabeth Nielsen, Jiaming Luo, Kat Black, Ryan Mullins, Sweta Agrawal, Wenda Xu, Erin Kats, Stephane Jaskiewicz, and 2 others. 2026. TranslateGemma technical report. *arXiv preprint arXiv:2601.09012*.

Markus Freitag, Nitika Mathur, Daniel Deutsch, Chi-Kiu Lo, Eleftherios Avramidis, Ricardo Rei, Brian Thompson, Frederic Blain, Tom Kocmi, Jiayi Wang, David Ifeoluwa Adelani, Marianna Buchicchio, Chrysoula Zerva, and Alon Lavie. 2024. Are LLMs breaking MT metrics? results of the WMT24 metrics shared task. In *Proceedings of the Ninth Conference on Machine Translation*, pages 47–81. Association for Computational Linguistics.

Gemma Team and Google DeepMind. 2026. Gemma 4 technical report. *arXiv preprint arXiv:2607.02770*.

GLM-5 Team. 2026. GLM-5: From vibe coding to agentic engineering. *arXiv preprint arXiv:2602.15763*.

Wenxiang Jiao, Jen-tse Huang, Wenxuan Wang, Zhiwei He, Tian Liang, Xing Wang, Shuming Shi, and Zhaopeng Tu. 2023. ParroT: Translating during chat using large language models tuned with human translation and feedback. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 15009–15020. Association for Computational Linguistics.

Juraj Juraska, Daniel Deutsch, Mara Finkelstein, and Markus Freitag. 2024. MetricX-24: The Google submission to the WMT 2024 metrics shared task. In *Proceedings of the Ninth Conference on Machine Translation*, pages 492–504. Association for Computational Linguistics.

Juraj Juraska, Tobias Domhan, Mara Finkelstein, Tetsuji Nakagawa, Geza Kovacs, Daniel Deutsch, Pidong Wang, and Markus Freitag. 2025. MetricX-25 and GemSpanEval: Google Translate submissions to the WMT25 evaluation shared task. In *Proceedings of the Tenth Conference on Machine Translation*, pages 957–968, Suzhou, China. Association for Computational Linguistics.

Juraj Juraska, Mara Finkelstein, Daniel Deutsch, Aditya Siddhant, Mehdi Mirzazadeh, and Markus Freitag. 2023. MetricX-23: The Google submission to the WMT 2023 metrics shared task. In *Proceedings of the Eighth Conference on Machine Translation*, pages 756–767. Association for Computational Linguistics.

Tom Kocmi, Ekaterina Artemova, Eleftherios Avramidis, Rachel Bawden, Ondřej Bojar, Konstantin Dranch, Anton Dvorkovich, Sergey Dukanov, Mark Fishel, Markus Freitag, Thamme Gowda, Roman Grundkiewicz, Barry Haddow, Marzena Karpinska, Philipp Koehn, Howard Lakougna, Jessica Lundin, Christof Monz, Kenton Murray, and 10 others. 2025. Findings of the WMT25 general machine translation shared task: Time to stop evaluating on easy test sets. In *Proceedings of the Tenth Conference on Machine Translation*, pages 355–413. Association for Computational Linguistics.

Tom Kocmi, Eleftherios Avramidis, Rachel Bawden, Ondřej Bojar, Anton Dvorkovich, Christian Federmann, Mark Fishel, Markus Freitag, Thamme Gowda, Roman Grundkiewicz, Barry Haddow, Marzena Karpinska, Philipp Koehn, Benjamin Marie, Christof Monz, Kenton Murray, Masaaki Nagata, Martin Popel, Maja Popović, and 3 others. 2024a. Findings of the WMT24 general machine translation shared task: The LLM era is here but MT is not solved yet. In *Proceedings of the Ninth Conference on Machine Translation*, pages 1–46. Association for Computational Linguistics.

Tom Kocmi and Christian Federmann. 2023. Large language models are state-of-the-art evaluators of translation quality. In *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, pages 193–203. European Association for Machine Translation.

Tom Kocmi, Vilém Zouhar, Eleftherios Avramidis, Roman Grundkiewicz, Marzena Karpinska, Maja Popović, Mrinmaya Sachan, and Mariya Shmatova. 2024b. Error span annotation: A balanced approach for human evaluation of machine translation. In *Proceedings of the Ninth Conference on Machine Translation*, pages 1440–1453. Association for Computational Linguistics.

Sneha Kudugunta, Isaac Caswell, Biao Zhang, Xavier Garcia, Derrick Xin, Aditya Kusupati, Romi Stella, Ankur Bapna, and Orhan Firat. 2023. MADLAD-400: A multilingual and document-level large audited dataset. In *Advances in Neural Information Processing Systems*, volume 36.

Alon Lavie, Greg Hanneman, Sweta Agrawal, Diptesh Kanojia, Chi-Kiu Lo, Vilém Zouhar, Frederic Blain, Chrysoula Zerva, Eleftherios Avramidis, Sourabh Deoghare, Archchana Sindhujan, Jiayi Wang, David Ifeoluwa Adelani, Brian Thompson, Tom Kocmi, Markus Freitag, and Daniel Deutsch. 2025. Findings of the WMT25 shared task on automated translation evaluation systems: Linguistic diversity is challenging and references still help. In *Proceedings of the Tenth Conference on Machine Translation*, pages 436–483, Suzhou, China. Association for Computational Linguistics.

Jiahuan Li, Hao Zhou, Shujian Huang, Shanbo Cheng, and Jiajun Chen. 2024. Eliciting the translation ability of large language models via multilingual finetuning with translation instructions. *Transactions of the Association for Computational Linguistics*, 12:576–592.

Oscar Li. 2024. Common crawl traditional chinese. Hugging Face dataset.

Yifeng Liu, Siqi Ouyang, Yatish H R, and Lei Li. 2026. Mending the holes: Mitigating reward hacking in reinforcement learning for multilingual translation. In *Findings of the Association for Computational Linguistics: ACL 2026*, pages 33699–33723. Association for Computational Linguistics.

Yingfeng Luo, Ziqiang Xu, Yuxuan Ouyang, Murun Yang, Dingyang Lin, Kaiyan Chang, Tong Zheng, Bei Li, Peinan Feng, Quan Du, Tong Xiao, and Jingbo Zhu. 2025. NiuTrans.LMT: Toward inclusive and scalable multilingual machine translation with LLMs. *arXiv preprint arXiv:2511.07003*.

MOP-LIWU Community and MNBVC Team. 2023. MNBVC: Massive never-ending BT vast chinese corpus. GitHub repository.

Guilherme Penedo, Hynek Kydlíček, Vinko Sabolčec, Bettina Messmer, Negar Foroutan, Amir Hossein Kargaran, Colin Raffel, Martin Jaggi, Leandro von Werra, and Thomas Wolf. 2025. FineWeb2: One pipeline to scale them all—adapting pre-training data processing to every language. *arXiv preprint arXiv:2506.20920*.

Qwen Team. 2026. Qwen3.5: Towards native multimodal agents.

Marc'Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. 2016. Sequence level training with recurrent neural networks. In *International Conference on Learning Representations*.

Ricardo Rei, Craig Stewart, Ana C. Farinha, and Alon Lavie. 2020. COMET: A neural framework for MT evaluation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pages 2685–2702. Association for Computational Linguistics.

Ricardo Rei, Marcos Treviso, Nuno M. Guerreiro, Chrysoula Zerva, Ana C. Farinha, Christine Maroti, José G. C. de Souza, Taisiya Glushkova, Duarte Alves, Luisa Coheur, Alon Lavie, and André F. T. Martins. 2022. CometKiwi: IST-unbabel 2022 submission for the quality estimation shared task. In *Proceedings of the Seventh Conference on Machine Translation*, pages 634–645. Association for Computational Linguistics.

Nathaniel Robinson, Perez Ogayo, David R. Mortensen, and Graham Neubig. 2023. ChatGPT MT: Competitive for high- (but not low-) resource languages. In *Proceedings of the Eighth Conference on Machine Translation*, pages 392–418. Association for Computational Linguistics.

Yuzhe Shang, Pengzhi Gao, Wei Liu, Jian Luan, and Jinsong Su. 2026. Scaling model and data for multilingual machine translation with open large language models. *arXiv preprint arXiv:2602.11961*.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, Y. K. Li, Y. Wu, and Daya Guo. 2024. DeepSeekMath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*.

Xiaofeng Shi, Lulu Zhao, Hua Zhou, and Donglin Hao. 2024. IndustryCorpus2. Hugging Face dataset.

Mingrui Sun, Mao Zheng, Zheng Li, and Mingyang Song. 2026. IFMTBench: A comprehensive benchmark for multilingual translation instruction following. *arXiv preprint arXiv:2605.28218*.

David Vilar, Markus Freitag, Colin Cherry, Jiaming Luo, Viresh Ratnakar, and George Foster. 2023. Prompting PaLM for translation: Assessing strategies and performance. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15406–15427. Association for Computational Linguistics.

Lijun Wu, Fei Tian, Tao Qin, Jianhuang Lai, and Tie-Yan Liu. 2018. A study of reinforcement learning for neural machine translation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3612–3621, Brussels, Belgium. Association for Computational Linguistics.

Haoran Xu, Young Jin Kim, Amr Sharaf, and Hany Hassan Awadalla. 2024a. A paradigm shift in machine translation: Boosting translation performance of large language models. In *The Twelfth International Conference on Learning Representations*.

Haoran Xu, Amr Sharaf, Yunmo Chen, Weiting Tan, Lingfeng Shen, Benjamin Van Durme, Kenton Murray, and Young Jin Kim. 2024b. Contrastive preference optimization: Pushing the boundaries of LLM performance in machine translation. In *Proceedings of the 41st International Conference on Machine Learning*, pages 55204–55224.

Brian Yan, Shuoyang Ding, Kuang-Da Wang, Siqi Ouyang, Oleksii Hrinchuk, Vitaly Lavrukhin, and Boris Ginsburg. 2025. Nvidia-Nemo's WMT 2025 metrics shared task submission. In *Proceedings of the Tenth Conference on Machine Translation*, pages 920–925, Suzhou, China. Association for Computational Linguistics.

Sen Yang, Shanbo Cheng, Lu Xu, Jianbing Zhang, and Shujian Huang. 2026. GRRM: Group relative reward modeling for machine translation. *arXiv preprint arXiv:2602.14028*.

Qiying Yu, Zheng Zhang, Ruofei Zhu, Yufeng Yuan, Xiaochen Zuo, Yu Yue, Tiantian Fan, Gaohong Liu, Lingjun Liu, Xin Liu, et al. 2025. DAPO: An open-source LLM reinforcement learning system at scale. *arXiv preprint arXiv:2503.14476*.

Yuze Zhao, Jintao Huang, Jinghan Hu, Xingjun Wang, Yunlin Mao, Daoze Zhang, Zeyinzi Jiang, Zhikai Wu, Baole Ai, Ang Wang, Wenmeng Zhou, and Yingda Chen. 2025. SWIFT: A scalable lightweight infrastructure for fine-tuning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 29733–29735.

Mao Zheng, Zheng Li, Tao Chen, Bo Lv, Mingrui Sun, Mingyang Song, Jinlong Song, Hong Huang, Decheng Wu, Hai Wang, Yifan Song, Yanfeng Chen, and Guanwei Zhang. 2026. Hy-MT2: A family of fast, efficient and powerful multilingual translation models in the wild. *arXiv preprint arXiv:2605.22064*.

Mao Zheng, Zheng Li, Yang Du, Bingxin Qu, and Mingyang Song. 2025. Shy-hunyuan-MT at WMT25 general machine translation shared task. In *Proceedings of the Tenth Conference on Machine Translation*, pages 607–613, Suzhou, China. Association for Computational Linguistics.

Maike Züfle, Vilém Zouhar, Tu Anh Dinh, Felipe Maia Polo, Jan Niehues, and Mrinmaya Sachan. 2025. COMET-poly: Machine translation metric grounded in other candidates. In *Proceedings of the Tenth Conference on Machine Translation*, pages 887–904, Suzhou, China. Association for Computational Linguistics.

A Additional Qualitative Cases

Figures 4–6 present additional instruction-following examples aligned with the requirement types in Table 1. As in Figure 3, each prompt lists the language pair, instruction, and source text; colored spans mark key constraints. Within a type, multiple contrastive cases are separated by dashed rules. HTML structure cases appear in Figure 3; the appendix adds JSON format control.

Type	Prompt	Output
Register / formality	*[English → Simplified Chinese]* **Instruction:** Render this informal complaint in a polished, professional customer-service register (non-vulgar) as natural Simplified Chinese. **Source:** what broadband do i get? ive had virgin for 8 years and holy shit is it slow for 40 quid monthly. Is it the old router or do i just switch to a cheaper bt or whatever	请问我该如何选择合适的宽带套餐？我已经使用了 Virgin 的服务八年，但每月 40 英镑的费用与目前的网络速度严重不符，速度实在太慢了。请问这是否是因为我的路由器型号过旧导致的？还是说，我应当考虑更换为 BT 或其他更实惠的运营商？
	[English → Simplified Chinese] **Instruction:** Preserve the informal, conversational, slightly rude register of this complaint in natural Simplified Chinese (no need to sanitize slang). **Source:** what broadband do i get? ive had virgin for 8 years and holy shit is it slow for 40 quid monthly. Is it the old router or do i just switch to a cheaper bt or whatever	我该办哪个宽带？用了8年Virgin了，真特么慢，一个月收我40镑。是路由器太老的问题，还是我直接换家更便宜的BT之类的？
Tone / emotion	*[English → Simplified Chinese]* **Instruction:** Translate as a warm airline customer-service notice with empathy and a brief sincere apology. Keep it concise: roughly the same information density as the source, without long boilerplate closings. **Source:** Flight AF123 is cancelled due to weather. You will be refunded within 7 days. Call the help center if you have questions.	由于天气原因，AF123航班已取消。我们深表歉意，您的退款将在7个工作日内办理。如有任何疑问，请随时联系帮助中心。
	[English → Simplified Chinese] **Instruction:** Translate as a cold, purely informational system notice with no apology, no empathy markers, and no softening language. **Source:** Flight AF123 is cancelled due to weather. You will be refunded within 7 days. Call the help center if you have questions.	航班AF123因天气原因取消。款项将在7日内退还。如有疑问，请联系帮助中心。
Translation strategy	*[English → Simplified Chinese]* **Instruction:** 归化翻译：把 401(k) 完全改写为中文养老金概念，译文中禁止出现拉丁字母或 401(k)。 **Source:** For many Americans, a 401(k) is the main way to save for retirement.	对于许多美国人来说，美国常见的由雇主发起的个人退休储蓄账户是积累退休金的主要方式。
	[English → Simplified Chinese] **Instruction:** Use a foreignization strategy. Keep '401(k)' visible as a U.S.-specific retirement savings plan; do not replace it with a localized Chinese product name. **Source:** For many Americans, a 401(k) is the main way to save for retirement.	对于许多美国人来说，401(k)是储蓄养老的主要方式。
Domain adaptation	*[English → Simplified Chinese]* **Instruction:** Translate for financial regulators into precise Simplified Chinese regulatory jargon; keep the formal technical nature with no simplification. **Source:** It is necessary to prevent regulatory arbitrage between the non-trading book and the trading book.	必须防止非交易账簿与交易账簿之间的监管套利。
Punctuation styling	*[English → Simplified Chinese]* **Instruction:** Preserve the expressive punctuation from the source; do not normalize it to ordinary single punctuation marks. **Source:** You're saying... we'll win the championship again after a five-year slump???!!! And then win it for six years straight after that!!!	你是说……我们会在经历五年的低谷期后再次赢得冠军？？？！！！并且从这之后连续赢了六年！！！

Figure 4: Additional cases (part 1): register/formality (professional vs. colloquial contrast), tone/emotion, translation strategy, domain adaptation, and punctuation styling.

Type	Prompt	Output
Fidelity–naturalness	*[English → Simplified Chinese]* **Instruction:** Prioritize naturalness: render "double-edged sword" as a natural abstract Chinese expression of pros and cons, not a sword metaphor. **Source:** AI programming is a double-edged sword for junior programmers.	AI 编程对初级程序员而言，利弊参半。
	[English → Simplified Chinese] **Instruction:** Prioritize fidelity: keep the "double-edged sword" metaphor in Chinese, rather than replacing it with a purely abstract paraphrase. **Source:** AI programming is a double-edged sword for junior programmers.	AI编程对于初级程序员来说是一把双刃剑。
Cultural localization	*[English → Simplified Chinese]* **Instruction:** Localize for a mainland Chinese social-media audience. Adapt the idiom 'Give yourself a pat on the back' into a natural Chinese expression (do not translate word-for-word as 拍背). Keep the high-energy fan tone. **Source:** Give yourself a pat on the back if you believed and trusted Tom from the start to the end without any unsureness and never lost hope in the series! Props to you!	从始至终都坚定地相信汤姆，从未有过犹豫，也从未对这部剧失去信心，真的为你点赞！太飒了！
	[English → Simplified Chinese] **Instruction:** Retain the British cultural reference 'Sky Q Mini Box' as-is (do not replace with a Chinese brand). Render final 'lol' as a resigned Chinese expression of dry amusement. **Source:** So I got a Sky Q Mini Box for upstairs. It doesn't work. I have to wait until late March for an engineer. lol	所以我给楼上弄了个 Sky Q Mini Box。结果根本不能用。我得等到三月底才有技术人员上门。真是绝了。
Audience adaptation	*[English → Simplified Chinese]* **Instruction:** 给小学生看：用最简单的词和短句，不要出现"叶绿素""光合作用""二氧化碳"。 **Source:** Through photosynthesis, plants produce their own food using sunlight. Chlorophyll inside the leaves captures light energy and helps convert water and carbon dioxide into sugar and oxygen.	植物利用阳光，自己做出食物。叶子能吸收阳光，把水和空气变成糖分和氧气。
	[English → Simplified Chinese] **Instruction:** 面向生物学专业读者，使用规范学术中文；必须使用：光合作用、叶绿素、二氧化碳。 **Source:** Through photosynthesis, plants produce their own food using sunlight. Chlorophyll inside the leaves captures light energy and helps convert water and carbon dioxide into sugar and oxygen.	植物通过光合作用利用阳光制造自身的养分。叶片中的叶绿素捕获光能，并帮助将水和二氧化碳转化为糖和氧气。
Glossary / terminology	*[English → Simplified Chinese]* **Instruction:** Apply glossary strictly: Cloud Native→云端原生化; Load Balancer→流量分发器; Container→隔离沙箱; Endpoint→服务触点; Traffic→数据洪流; Secure Protocol→高保密协定. **Source:** In a Cloud Native architecture, the Load Balancer routes the Traffic to each Container using a Secure Protocol via the designated Endpoint.	在云端原生化架构中，流量分发器通过指定的服务触点，利用高保密协定将数据洪流路由到每个隔离沙箱。
	[English → Simplified Chinese] **Instruction:** Apply glossary strictly: Cloud Native→云原生; Load Balancer→负载均衡器; Container→容器; Endpoint→端点; Traffic→流量; Secure Protocol→安全协议. **Source:** In a Cloud Native architecture, the Load Balancer routes the Traffic to each Container using a Secure Protocol via the designated Endpoint.	在云原生架构中，负载均衡器通过指定的端点，使用安全协议将流量路由到每个容器。

Figure 5: Additional cases (part 2): fidelity–naturalness trade-off, cultural localization, audience adaptation, and glossary/terminology.

Type	Prompt	Output
Purpose / scenario	*[English → Simplified Chinese]* **Instruction:** Scenario: two coworkers chatting privately after an all-hands. Translate as natural, slightly cynical workplace slang in Simplified Chinese. **Source:** We are currently implementing a strategic organizational restructuring to streamline operations and empower our remaining workforce with more dynamic and expansive roles.	“公司现在正在搞‘战略性组织架构调整’，美其名曰优化流程，给剩下的员工分派更‘动态’、更‘广泛’的职责。”
	[English → Simplified Chinese] **Instruction:** Scenario: external corporate PR release. Translate into polished, formal Simplified Chinese business language suitable for public statement. **Source:** We are currently implementing a strategic organizational restructuring to streamline operations and empower our remaining workforce with more dynamic and expansive roles.	我们目前正在实施一项战略性组织重组，旨在优化运营流程，并为现有员工赋予更具活力且更广泛的职责。
Implicitness preservation	*[English → Simplified Chinese]* **Instruction:** Preserve informal register in natural Simplified Chinese. Translate all English slang including ’imo’ and ’kinda neat’ into casual Chinese equivalents. Do NOT leave English words/phrases in the output, and do NOT add explanatory glosses or parenthetical definitions. **Source:** I’m not a trans man, but I am interested in knowing someone’s perspectives who lived both as a woman and as a man (not a lot of people get to experience what it is like to be perceived as both genders) so thats kinda neat imo	我不是跨性别男性，但我挺好奇那些既以女性身份又以男性身份生活过的人是怎么看的（没多少人能体验过被外界视为两种不同性别的感觉），所以我觉得这还挺酷的。
	[Japanese → English] **Instruction:** Translate this Japanese dialogue into English subtitles. Keep honorific suffixes (-san, -kun, senpai) untranslated in the English. **Source:** たけし：「先輩、これは当たり前に大丈夫なんですか？」ゆうこ先輩：「だからさっきから言ってるじゃん。心配しすぎだよ、たけし君。」	Takeshi: ”Senpai, is this really okay as it is?” Yuko Senpai: ”I’ve been saying that since a while ago, haven’t I? You’re worrying too much, Takeshi-kun.”
Verbosity preference	*[English → Simplified Chinese]* **Instruction:** Translate into extremely concise, telegraphic Simplified Chinese suitable for a match ticker. No complete-sentence padding. **Source:** Most junior black belts are around 14-18 years of age and have been practising jujitsu for more than 6 or 7 years.	多数黑带青少年约14-18岁，习练柔术6-7年以上。
	[English → Simplified Chinese] **Instruction:** Translate into complete, smoothly flowing Simplified Chinese sentences; you may mildly expand for readability but must not add new facts. **Source:** Most junior black belts are around 14-18 years of age and have been practising jujitsu for more than 6 or 7 years.	大多数初级黑带学员的年龄在14到18岁之间，练习柔术的时间已经超过了6或7年。
Additional context	*[English → Simplified Chinese]* **Instruction:** Additional context: a baseball radio commentator describing the previous play. **Source:** The pitch was perfect, and nobody saw it coming.	这球投得简直完美，完全没人料到！
	[English → Simplified Chinese] **Instruction:** Additional context: colleagues debriefing after a startup sales presentation. **Source:** The pitch was perfect, and nobody saw it coming.	演示得完美无缺，完全出乎大家的意料。
Structure / format	*[English → Simplified Chinese]* **Instruction:** Translate JSON string values into Simplified Chinese; keep all keys in English; valid JSON only. **Source:** `{"title":"Freshman Welcome Week","activity":"Dormitory Tour","location":"Zijing Apartment Building 2, Room 524A","detail":"New students can visit a sample dorm room and meet peer mentors.","status":"open"}`	{"title":"新生欢迎周","activity":"宿舍参观","location":"紫荆公寓2号楼524A","detail":"新生可以参观样板宿舍并结识学长学姐。","status":"开放"}
	[English → Simplified Chinese] **Instruction:** Translate both JSON keys and values into Simplified Chinese; do not leave English keys; valid JSON only. **Source:** `{"title":"Freshman Welcome Week","activity":"Dormitory Tour","location":"Zijing Apartment Building 2, Room 524A","detail":"New students can visit a sample dorm room and meet peer mentors.","status":"open"}`	{"标题":"新生欢迎周","活动":"宿舍参观","地点":"紫荆公寓2号楼524A","详情":"新生可以参观样板宿舍并结识学长学姐。","状态":"开放"}

Figure 6: Additional cases (part 3): purpose/scenario, implicitness preservation, verbosity preference, additional context, and structure/format (JSON).