

Cross-Lingual Legal QA for Vietnamese Labour Law: Retrieval, Translation, and Verifier-Guided Correction

Nguyen Minh Chi¹, Mo El-Haj^{1,3}, Nguyen Ha Thanh¹, Dawn Knight² and Paul Rayson^{3,1}

¹VinUniversity, Vietnam ²Cardiff University, UK ³Lancaster University, UK

Abstract

Cross-lingual legal question answering must retrieve statutes across languages while preventing unsupported legal claims. We introduce a bilingual evaluation suite of 231 Vietnamese–English question–answer pairs from Vietnamese labour law. Of these, 75 are additionally annotated for five challenging legal reasoning phenomena. We evaluate a verifier-guided pipeline that decomposes answers into claims, checks citation reachability and entailment, and corrects citation failures and contradictions. We also introduce six automatic diagnostics for faithfulness to retrieved evidence, covering citations, modality, exceptions, procedures, conclusions, and evidential support. Experiments show that learned-sparse retrieval performs poorly for English-to-Vietnamese retrieval ($R@5 = 0.032$), whereas dense retrieval reaches 0.358 and slightly outperforms hybrid retrieval. Translation placement has no statistically detectable effect on these automatic diagnostics in our controlled comparison and supporting sensitivity analyses. Verifier-guided correction improves citation preservation by 0.022–0.034 at the system level but produces no reliable gains in the remaining dimensions. Human evaluation further shows that the automatic diagnostics do not fully align with human judgements of answer quality.