

The datasets of human and Ai translation

ABSTRACT

The datasets outline a methodical approach for comparing translations from Mandarin to Malay completed by humans and AI. The datasets contain a framework with rubrics for a keyword detection template to discover common words used by both humans and AI in their translation tasks. The datasets include a Mandarin poem originally written in Mandarin, along with translations by a belt-and-road project translator, a Mandarin native speaker translator, and a Malay native speaker translator categorised as human translators. It also contains translations by ChatGPT 3.5 with 1 prompt, and ChatGPT 4.0 with four different prompts. All these prompts are also listed in the dataset.