

ProTranslator EXPRESS: 生成AIとドメイン適応型機械翻訳を統合化した翻訳支援ツール

LLM(大規模言語モデル)による以下の機能により、CATツールはLLM Assisted Translationへと進化します。

1. smart-TermBase® (OpenAI社API使用)

プロジェクトのターム・ベースの訳文への訳語適用 (冠詞や単複表現、動詞+副詞など語尾変化への訳語自動調整) をLLM (gpt-4)が行います。

特願2023-106225

2. smart-Fixer™ (OpenAI社API使用)

LLMにより、OCRやpdfなどの原稿のスペルミスやワード文書のタイプミスの文脈に則した修正やスペース欠損による単語連結の修正、誤字・脱字修正、pdf文書の頁/段組み境界での文の分断の修正、テキストボックス折り返し部での文の分断の修正、合字(Ligature)修正、不要なハイフネーション修正さらに新規略字によるセグメンテーション誤りや複数文が1つのセグメントに入るセグメンテーション誤りの自動検知とその修正などをLLM (gpt-4)が行います。さらに特許明細書の明晰化も可能です。smart-Fixer™は、生成AIオプションのお客様でご利用いただけます。

特願2023-200719

3. smart-TranslationMemory™ (OpenAI社API使用)

これまでTM閾値以下のファジーマッチセグメントは、MT訳を使用していました。smart-TranslationMemory™により、65%以上のマッチ率セグメントについては、LLMを用いた高精度smart-TranslationMemory™が挿入されます。smart-TranslationMemory™では、LLMを用いたファジーマッチTMによるその場アダプテーションで、弊社評価結果ではMTアダプテーションより高精度な翻訳が可能となります。**プライマリメモリを用いて、翻訳者自身が確定させた翻訳結果を未確定セグメントに反映させることも可能です。**smart-TranslationMemory™を用いると、TM閾値以上ではTMを、65%~TM閾値までのセグメントはsmart-TranslationMemory™を、65%未満のセグメントは指定したMTエンジン(含むマルチNMT)によるMT訳が挿入されます。これまでLLM翻訳で難しかったインラインタグの挿入もLLMが行います。smart-TranslationMemory™は、生成AIオプション契約のお客様でご利用いただけます。

特願2024-24059

	実施例1	実施例2	比較例1	比較例2	比較例3	比較例4	比較例5	比較例6
G06F特許請求項 英日翻訳	smart- TranslationMe mory™	smart- TranslationMe mory™	gpt3.5	gpt3.5 fine_tuningのみ	インコンテキスト 法	memoQ AGT domain adaptation 最小マッチ率>65%	NICT特許NT	NICTアダプテー ション
N(サンプル数)	1	50	50	50	50	45	50	50
BLEU	77	70	34	46	46	48	57	62
RIBES	91	85	57	84	62	63	88	91
スコア	81	74	41	58	50	52	66	71

4. Rsource-Generator (OpenAI社API使用)

原文ワードファイルとその人手翻訳ワードファイルのペア・ファイルがあれば、それらをdrag&dropするだけで翻訳メモリと用語ベースのインポート用エクセルファイルが生成します。従来突合せツール+専門家による選別+修正作業をLLMが代行するため、抽出に関する人手作業が不要になります。(最終選別作業は必要)

Rsource-Generatorは、生成AIオプション契約のお客様でご利用いただけます。

特願2024-033396

5. LLMアダプテーション (OpenAI社API使用)

対訳ファイルがあれば、LLMアダプテーションエンジン(LLMとしてgpt-3.5 fine_tuningを使用)を構築することが可能です。NICTアダプテーションエンジンもこれまで通りご利用可能です。目安として、5,000文対以上であればNICTを、それ以下であればLLMアダプテーションをご利用ください。LLMアダプテーションでは、NICTへの対訳提供は不要です。LLMアダプテーションは、生成AIオプション+アダプテーションエンジン契約のお客様でご利用いただけます。

特願2024-24059