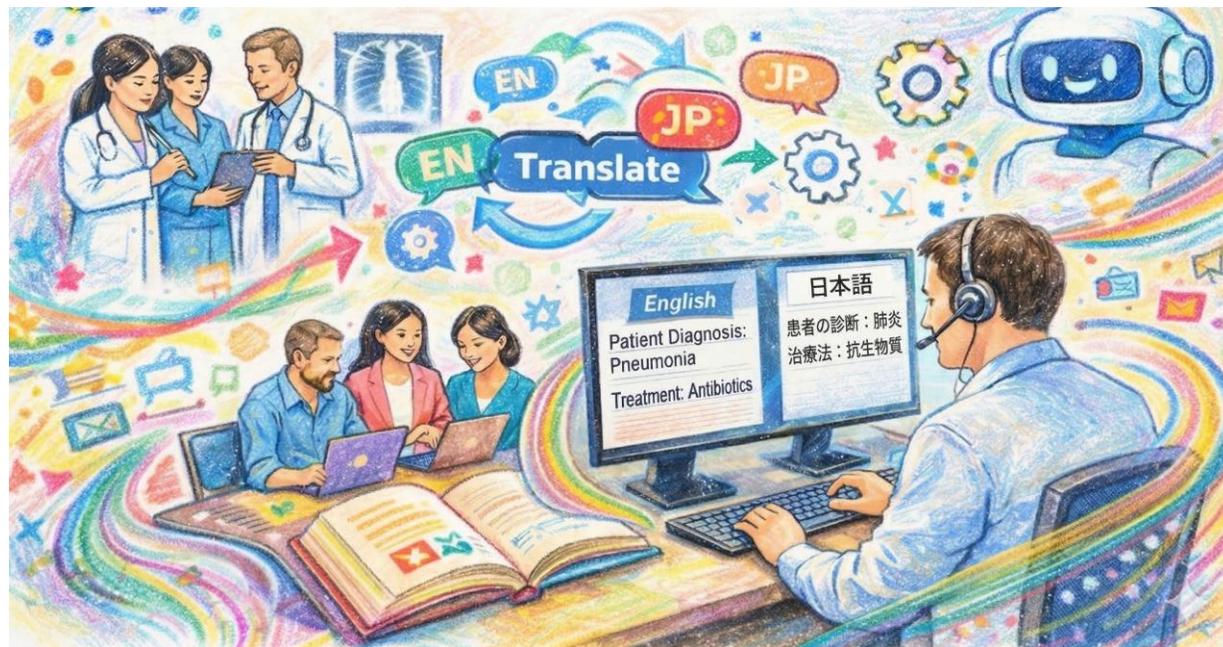


医療分野のAI翻訳システムにおけるHuman-in-the-Loop： 翻訳者パネル構想



イラスト：GPT-5.3で作成 + Nano Banana 2で修正

ローズ三浦 / エッジ・トランスレーション

自己紹介



ローズ三浦

医療翻訳者/翻訳エージェント。翻訳経験は約13年。
AAMT個人会員。MTユーザーガイド委員。
人間の翻訳者とAI翻訳の共生を模索している。

自身で生成AIを活用した医療特化AI翻訳システム「生成AI
ローズ三浦モデル」を開発した。

<https://www.genai-rose-miura-model.jp/>

直近の発表：AAMT 2025, Tokyo 「生成AIの時代に翻訳者が
主導する新しい翻訳システム」（津山逸と共同のポスター
発表）

NLP年次大会シルバースポンサー：2019～2026年（2020
年を除く）

1. 医療分野のAI翻訳の品質保証の限界

- ・ AI翻訳単独では品質保証が不完全である
- ・ AI（NMTやLLM）は、Transformerによって
次の単語を予測しているだけ

（※NMTはEncoder + Decoder、LLMはDecoderのみなど、細かい部分は違いあり）

- ・ LLMには**ハルシネーション**等、特有の問題もある

→ この問題を解決するためのHuman-in-the-Loopとして、人間の翻訳者が「**監訳**」（**ポストエディット**）を行う

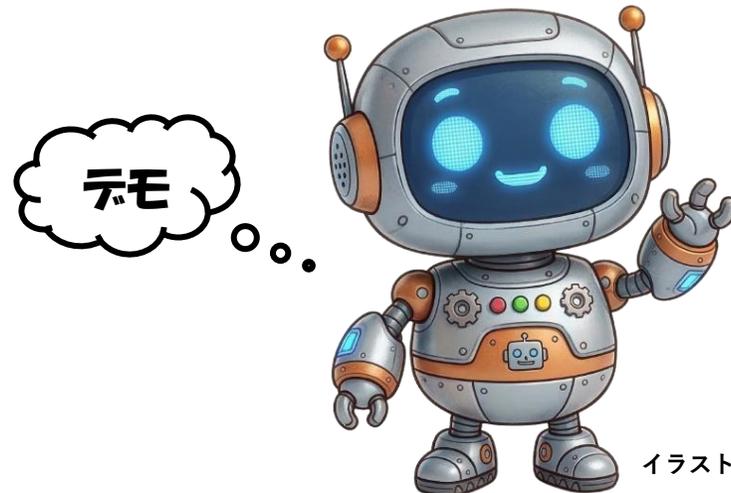
※**監訳**：ポストエディットの新しい呼び方（AAMT 2025, Tokyoにて津山逸氏が提唱）

2. 翻訳者パネルの設計思想

- ・ AI翻訳の出力を人間の翻訳者が**監訳**
- ・ 翻訳者に新しいスタイルの仕事を提供：

例：Tierによって報酬が変動する、貢献度によってボーナスが出る、etc.

- ・ 実務データから研究データも同時取得※**実務案件は要契約**
- ・ 研修・トレーニングも同パネル上で提供



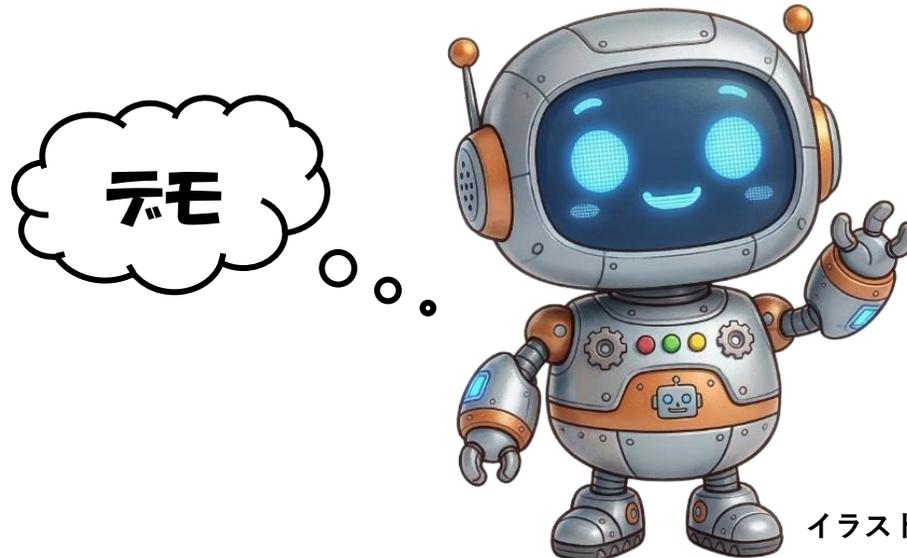
イラスト：Nano Banana 2で作成

3. 「生成AIローズ三浦モデル」におけるHuman-in-the-Loopの実験

- ・ 高レベルの翻訳者の監訳・フィードバックを
エージェント機能で自動収集

→ AI翻訳に**自動で学習され、次回から反映される**

※実務案件は要契約



イラスト：Nano Banana 2で作成

3. 「生成AIローズ三浦モデル」におけるHuman-in-the-Loopの実験

品質指標抽出サブエージェントを作成

- 品質指標抽出サブエージェントを作成しました。

作成したファイル

1. メインサブエージェント

src/subagents/qualityMetricsExtractor.ts

- AI学習可能な品質指標を抽出するコアロジック
- 11種類のエラーカテゴリ分類 (terminology, grammar, style, omission等)
- 4段階の重大度 (critical, major, minor, preferential)
- 文レベル・文書レベルの品質指標
- 学習可能シグナル (LearnableSignal) の生成

2. APIエンドポイント

src/app/api/analyze/quality-metrics/route.ts

- GET: 全翻訳データの品質指標を取得
- POST: 単一翻訳の即時分析
- ?format=jsonl: AI学習用JSON Lines形式でエクスポート

3. CLISクリプト

scripts/run-quality-extractor.ts

npx ts-node scripts/run-quality-extractor.ts --format jsonl --output data/training.jsonl

4. Reactフック

src/hooks/useQualityMetrics.ts

- fetchAllMetrics(): 全データ取得
- analyzeTranslation(): 単一分析
- getTopErrors(): 上位エラーカテゴリ
- getQualityTrend(): 日付別品質推移

抽出される主な品質指標

指標	説明
editRate	編集率 (0-1)
learnableSignals	AI学習用シグナル配列
categoryBreakdown	エラーカテゴリ別集計
learningTarget	学習ターゲット (vocabulary/grammar/style/domain/context)
topCorrections	頻出修正パターン

原文: This is a multicenter, randomized, double-blind, placebo-controlled clinical trial. The primary endpoint is to evaluate efficacy and safety.

AI翻訳: これは多施設共同、無作為化、二重盲検、プラセボ対照試験である。主要評価項目は有効性および安全性を評価することである。

編集後翻訳: 本試験は多施設共同、無作為化、二重盲検、プラセボ対照試験である。主要評価項目は有効性および安全性を評価することである。

AI翻訳「これ」
→ 翻訳者「本試験」に監訳

差分分析結果

- 編集率: 8.6%
- 品質スコア: 87点
- 変更箇所: +3 / -2

差分詳細

***☆良好

軽微な修正のみで使用可能な品質でした。

1件の改善ルールを学習しました

完了 (新規タスクへ)

「タスクを提出」
ボタンを押すと、
差分分析結果が
AI翻訳に送られる

原文: This is a multicenter, randomized, double-blind, placebo-controlled clinical trial. The primary endpoint is to evaluate efficacy and safety.

AI翻訳: 本試験は多施設共同、無作為化、二重盲検、プラセボ対照試験である。主要評価項目は有効性および安全性を評価することである。

編集後翻訳: 本試験は多施設共同、無作為化、二重盲検、プラセボ対照試験である。主要評価項目は有効性および安全性を評価することである。

「品質指標抽出サブエージェント」によってAI翻訳に自動で学習され、
次回から反映される

※実験はサンプルデータで実施。実務案件はクライアント・翻訳者と要契約

次世代型翻訳システム：『持続可能な翻訳インフラ』としての役割



図：Nano Banana 2で作成

4. 今後の展開

AI Makes Translators Great Again

翻訳者からのフィードバックを積極的に反映したAI翻訳が市場シェアを取れば、

翻訳者の訳語が日本語として定着する(?)

訳語や表現は翻訳者が決める。

翻訳者が自ら協力することによって、

翻訳者は明治時代のような翻訳の『主権』を取り戻すことができる。

Discord 「生成AIに向き合う翻訳者の会」

- 「生成AIに向き合う翻訳者の会」
<https://discord.gg/DZzHW8mNaH>

- ☆ 無料（Discordへの無料登録が必要です）
- ☆ 生成AIローズ三浦モデルの製品版を公開済み（割引コードで無料）
- ☆ 翻訳者パネルのプロトタイプを公開済み
- ☆ ディスカッションや意見交換の場も設けています

ぜひ参加してね！

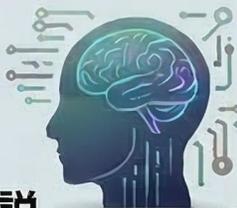


まとめ

医療分野のAI翻訳システムにおけるHuman-in-the-Loop：翻訳者パネル構想

1. 医療分野のAI翻訳の品質保証の限界

AI翻訳単独では品質保証が不完全である。
この問題を解決するためのHuman-in-the-Loopの重要性を解説



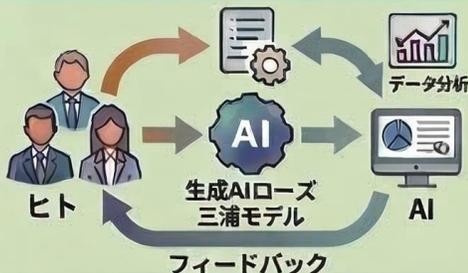
2. 翻訳者パネルの設計思想



- 各専門分野の翻訳家が監訳
- Tierによって報酬が変動+ボーナス
- 実務データから研究データも同時取得
- 研修・トレーニングも同パネル上で提供

翻訳者が標準を創る時代へ

3. 「生成AIローズ三浦モデル」におけるHuman-in-the-Loopの実験



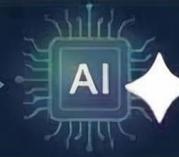
高レベルの翻訳者の監訳・フィードバックをエージェント機能で自動収集→システムに自動反映

4. 今後の展開：

AI Makes Translators Great Again

翻訳者からのフィードバックを積極的に反映したAI翻訳が市場シェアを取れば、翻訳者の訳語が日本語として定着する(?) 翻訳者が自ら協力することにより、翻訳者は明治時代のような「翻訳の『主権』」を取り戻すことができる。

翻訳者の暗黙知を、AIが学習可能な品質指標として顕在化させる



では、始めましょう！