



Roche ロシュ グループ

最適化された機械翻訳導入モデルによる 医薬品開発業務の迅速化への試み

中外製薬株式会社

信頼性保証企画部 翻訳マネジメントグループ

鞠子幸泰

株式会社アスカコーポレーション

業務推進部

早川威士

2021年12月9日 (木) 公募セッション(4) 13:00-15:00



医薬品開発業務への機械翻訳（MT）導入

- **AI技術への期待**

- 人工知能(AI)は産業界にとっても強力なツール
- しかしAIは曖昧さはらむ技術でもある
 - 目的と期待することを明確にしなければ導入効果を可視化できない

- **MTの導入**

- 医薬品業界のグローバル展開に伴い、翻訳業務は避けて通れない
- MTの導入に何を期待するのか
 - 人手翻訳と比べて品質を落とすことなく、さらなる翻訳スピードを追求すること

MTへの期待と役割

• MTのメリットと限界

- 翻訳は本来は時間がかかる仕事であり、ボトルネックでもあった
- MTの処理は一瞬だが、そのままでは使えない
 - 業務で使う翻訳には品質の保証が必要
 - ポストエディット(PE)などの品質管理タスクを入れるのが前提

• 翻訳時間の最適化

- 品質を確保した上で、PEを含めた全体のプロセスの迅速化を考える必要がある
- PEにかかる時間の最適化
 - MT出力の品質を高くする
 - PE作業者のこだわり排除や依頼者要求を超えない→品質水準の設定

MT導入モデルの開発

• 取り組みの概要

- 翻訳精度の高いMTの開発と評価
 - 自社の翻訳データをパラレルコーパスとして、MTをカスタマイズする
 - カスタマイズMTの品質を評価
 - PEの作業時間を短縮できるかをも考慮した評価設計
- PE時間短縮の実証的研究
 - PEを同条件で人手翻訳と比較して時間短縮効果を見る
 - 開発したカスタマイズMTを使用
 - 過剰な作業をさせないため事前に品質水準を設計

開発したカスタマイズMT

• MTエンジンの開発

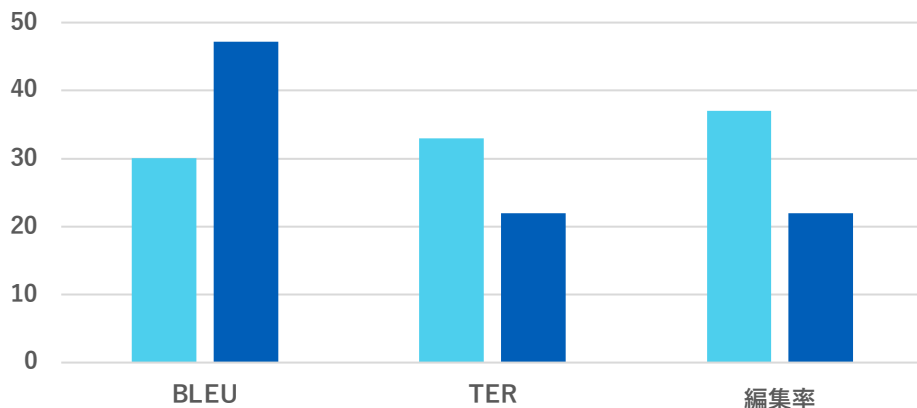
- 既存の基盤エンジンに対し、自社の翻訳文書24万文を含むデータで転移学習（domain adaptation）を行い、カスタマイズMTモデルを作成

• MTの評価指標

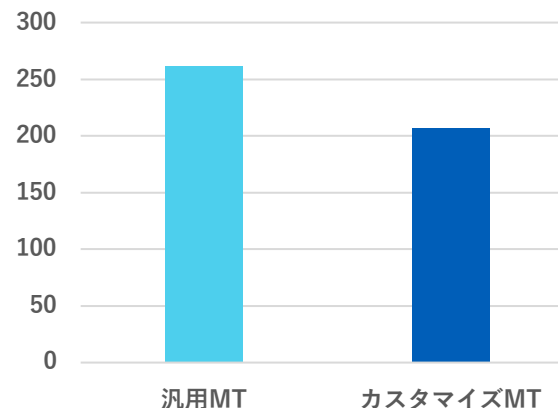
- BLEU、TER (translation error ratio)
 - 広く用いられている自動評価指標で基本的な性能を確認
- PEの編集率、編集時間
 - PEの実務的効率（作業量、作業時間短縮）を評価
- 固有表現一致度
 - 翻訳品質に影響する見出し語や専門用語など、用語表現の均一性を評価

カスタマイズMTの評価

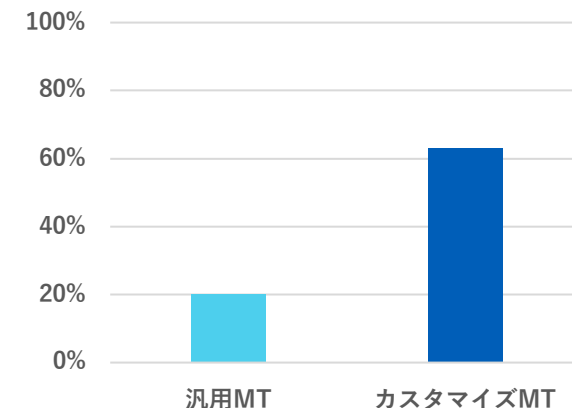
BLEU, TER, 編集率 (英日) (%) n=220(文)



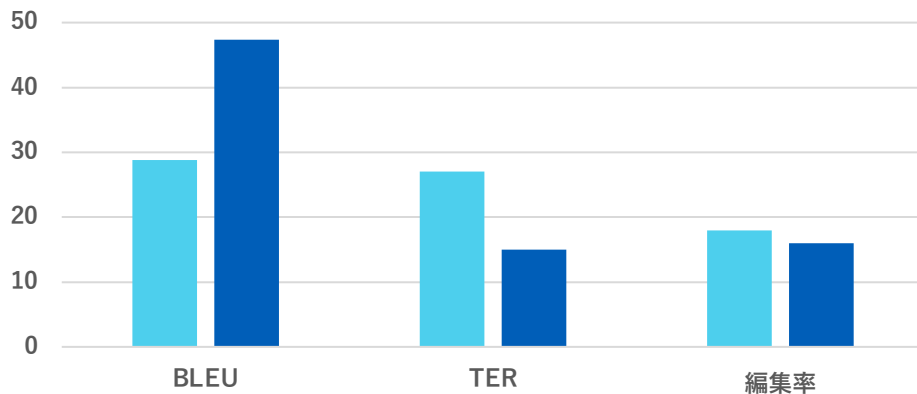
編集時間 (英日) (秒) n=220(文)



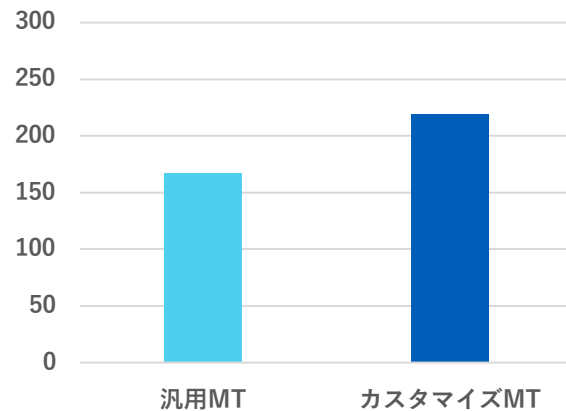
固有表現一致度 (英日) n=220(文)



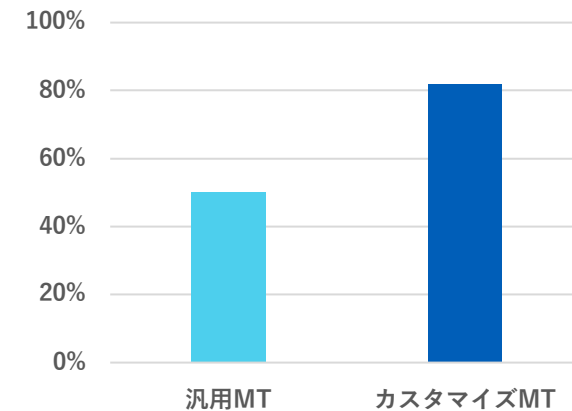
BLEU, TER, 編集率 (日英) (%) n=220(文)



編集時間 (日英) (秒) n=220(文)



固有表現一致度 (日英) n=220(文)



■ 汎用MT ■ カスタマイズMT

品質水準の策定

- **効率的PEに繋がる事前の品質水準設定**

- 期待する品質をポストエディターと共有することで、品質水準以上の余計な作業を避ける
- 作業上での判断の迷いが時間のかかる要因であることがわかったため、事前に定義しておき思考負荷を下げる

- **品質水準の内容**

- 要求事項と非要求事項をリストアップする
 - やらなくていいこと・やらないほうがいいことの明確化
- 要求事項：正確性や可読性など優先度が高い項目
- 非要求事項：表現の好みや一貫性など

実証的研究の実施

• 研究のデザイン

- 治験実施計画書（43,000語相当）の人手翻訳とMT+PEにかかる時間を比較
- 2回行い、MT+PE時間短縮効果に再現性があるかを確認
- 1回目で検出された品質の問題点に対して2回目では対策を行い、改善効果を確認

• 目標

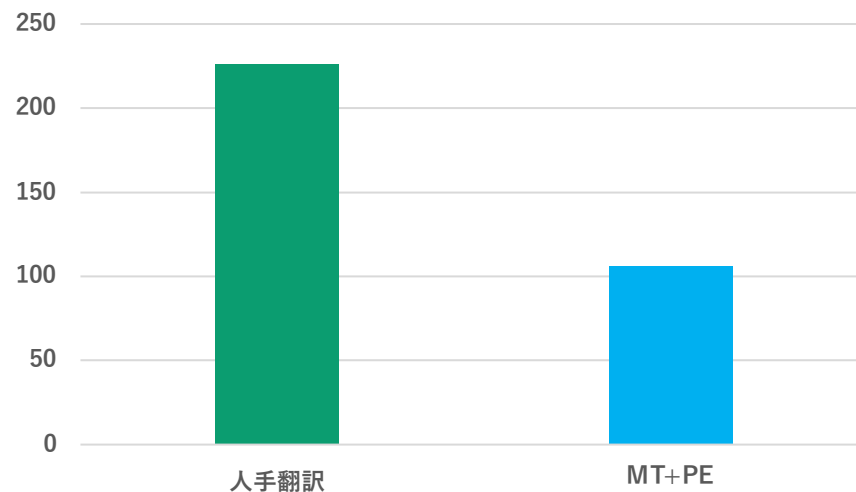
- 人手翻訳に対しPEで30%以上の時間短縮効果がみられること
- 設定した品質水準が達成されていること

実証的研究の結果

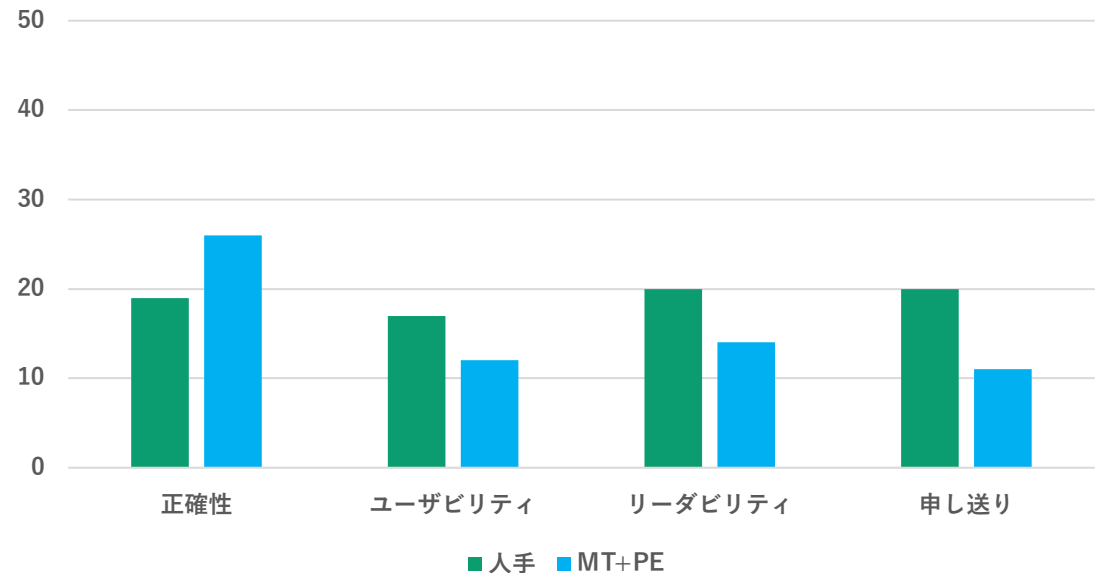
• 1回目

- 人手翻訳に対し、PEでは48%の時間短縮効果を確認
- 一方で品質についてはエラーの多さに課題があり、2回目ではエラーを削減して品質水準に近づけられるかどうか併せて確認することに

治験実施計画書の翻訳に要した時間 (h)



エラー数の比較(1回目) n=2,398(文)/26,612(語)

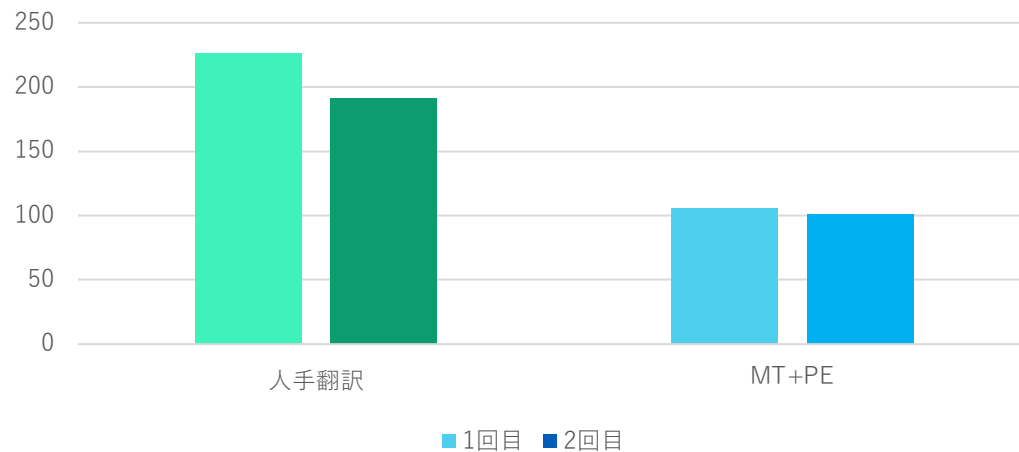


実証的研究の結果

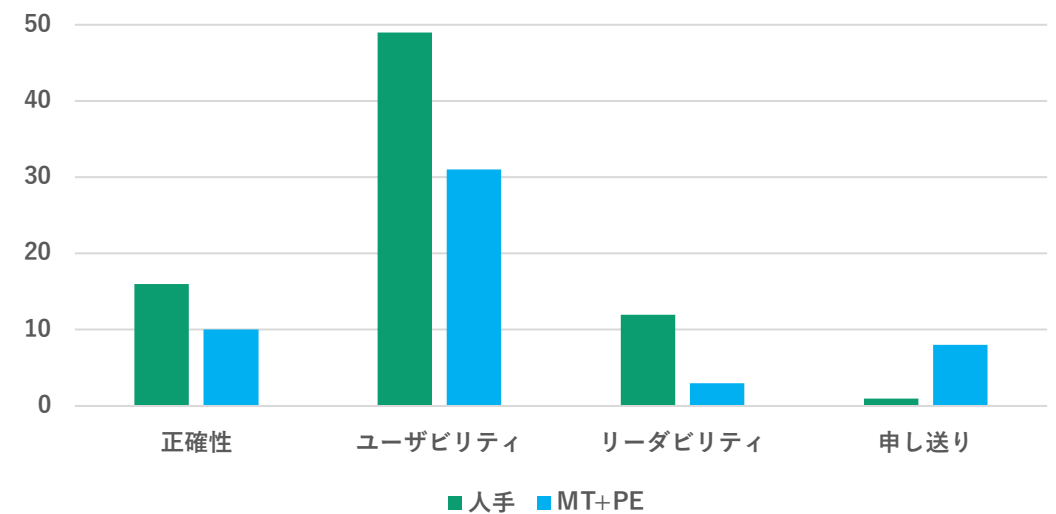
• 2回目

- 1回目に引き続き、人手翻訳に対してPEで43%の時間短縮効果を確認。
実時間も1回目から短縮された
- 品質についても1回目よりエラーが削減されていることが確認された
 - 作業指示書の更新やフィードバックなど、ポストエディターとのコミュニケーションが品質改善につながる

治験実施計画書の翻訳に要した時間 (h)



エラー数の比較(2回目) n=2,101(文)/29,118(語)



• 科学的手法とエビデンスに基づくMTの導入プロセス

- 目的として、人手翻訳と比べて品質を落とすことなく、さらなる翻訳スピードを追求することを設定し、MTの評価を試みた
- 時間短縮効果に影響するパラメータの仮説とその検証を実施した
 - ✓ PE時間を短縮できるようなMTの精度改善
- 実測データによる比較を行い、MT導入による改善を確認できた
 - ✓ ポストエディットの定量的データを取得
- 今後は、実装における実測データをもとに、エビデンスの再現性を確かめる必要あり

• 翻訳品質とスピードの最適化

- 品質水準を明示的に文書化することにより、品質を保持しつつ迅速化を達成できることが確認された