

AAMT2022公募セッション

海外発の新技术を題材にした機械翻訳の「滑らか度」の考察

株式会社ファイブアイズ・イングリッシュ 所属
太田 圭佑

事例発表の概要

コネクテッドカー業界における表現の機械翻訳の対応状況を題材に、

翻訳エンジンにおける**新しいビジネスへの対応状況**について考える。

発表者プロフィール

太田 圭佑

フリーランス産業翻訳家（日英西語）

金融・IT分野トライアル合格（英日）

＜翻訳実績＞

- 厚生労働省の新型コロナウイルス感染症関連
- 東京2020オリンピックパラリンピック関連
- ESG投資関連
- コネクテッドカー業界関連
- 幼児向け絵本など

前提

- **世界最先端のビジネス**は海外（英語圏）発祥
→ 数年遅れて日本に上陸する。
(例) NFT、DAO、Web3.0など...
- 当然、当初はほとんどの文章は英語で書かれている。

前提

- ・ 日本に新しいビジネス／概念が到達した時点では、新しい単語の訳が定まっていない。

→一般名詞化していない→学習データが少ない。

(例) 例えば、「バーンレート」の場合、「burn rate」をGoogle翻訳すると「燃焼速度」と訳される。

- ・ ある程度、日本の市場に浸透した段階で、統一された訳語が定まり、一般名詞化される。

仮説

- ・ 新しい用語を使用した英文が日本語に訳されるようになると、**英→日のディープラーニングを通じて、AIが学習し、対応できるようになる。**
- ・ 逆にいうと、**AIが新しいビジネスの英文を正しく訳せるかどうかで、日本にその新しいビジネスがどれくらい浸透しているかを測ることができるのでは？**

 過去の翻訳案件を題材に、この仮説について考える

題材について

- ・ **コネクテッドカー**における**テレマティクス技術**※1をベースにしたソリューションを提供する海外の会社のホワイトペーパー、プレスリリースなどを翻訳（2020年）

※1 テレコミュニケーション(通信)とインフォマティクス(情報処理)を組み合わせた造語。通信を利用してクルマにさまざまなサービスを提供する自動車用情報通信技術の総称

専門用語

英語	日本語
Tire 1 supplier	Tier1 サプライヤー
incident	インシデント
connected services	コネクテッドサービス
connected car	コネクテッドカー
data breach	データ侵害
DMV	DMV（米国車両管理局）

誤訳の例

英語	日本語
Tire 1 supplier	一次部品メーカー
incident	事象、出来事
connected services	接続されたサービス
connected car	接続した車両
data breach	データ侵害
DMV	DMV（米国車両管理局）

2020年と2022年の翻訳結果の比較①

【原文】

Certain details of the incident are publicly available in (サービス名1) , with a full complete analysis available to (サービス名2) customers.

【2020 Google翻訳】

出来事の確かな詳細は (サービス名2) の顧客への全て完全な解析結果が利用できるのと同様に (サービス名1) の中で公に利用することができます。

2020年と2022年の翻訳結果の比較①

【原文】

Certain details of the incident are publicly available in (サービス名1), with a full complete analysis available to (サービス名2) customers.

【2022 Google翻訳】

インシデントの特定の詳細は(サービス名1)で公開されており、(サービス名2)のお客様は完全な分析を利用できます。

2020年と2022年の翻訳結果の比較①

【2020 Google翻訳】

出来事の確かな詳細は（サービス名2）の顧客への全て完全な解析結果が利用できるのと同様に（サービス名1）の中で公に利用することができます。

【2022 Google翻訳】

インシデントの特定の詳細は（サービス名1）で公開されており、（サービス名2）のお客様は完全な分析を利用できます。

2020年と2022年の翻訳結果の比較②

【原文】

One of the most formidable aspects of a connected car is the potential for a hacker to manipulate or control the car's functions.

【2020 Google翻訳】

接続された車の最も手ごわい側面の1つは、ハッカーが車の機能を操作または制御する可能性です。

2020年と2022年の翻訳結果の比較②

【原文】

One of the most formidable aspects of a connected car is the potential for a hacker to manipulate or control the car's functions.

【2022 Google 翻訳】

コネクテッドカーの最も手ごわい側面の1つは、ハッカーが車の機能を操作または制御する可能性です。

2020年と2022年の翻訳結果の比較②

【2020 Google翻訳】

接続された車の最も手ごわい側面の1つは、ハッカーが車の機能进行操作または制御する可能性です。

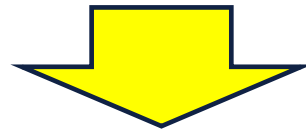
【2022 Google翻訳】

コネクテッドカーの最も手ごわい側面の1つは、ハッカーが車の機能进行操作または制御する可能性です。

コネクテッドカー業界の進化

【2015年】コネクテッドカーの普及とともに、テレマティクス自動車保険がに登場

【2020年】**道路交通法が一部改正**、**あおり運転に対する罰則が創設**



安全運転に対する意識が高まり

ドライバーごとに保険料が細かく算定される

テレマティクス自動車保険の普及が進む

コネクテッドカー業界の進化

- 2021年度の国内個人向けテレマティクス保険市場は約2,260億円
- 2021年度から2025年度までのCAGR（年平均成長率）は14.6%で成長していく見通し

➡ 日本市場への浸透とともに翻訳の精度が上がる仮説は正しそう

仮説を前提にした考察

日本市場への浸透とともに翻訳の精度が上がる

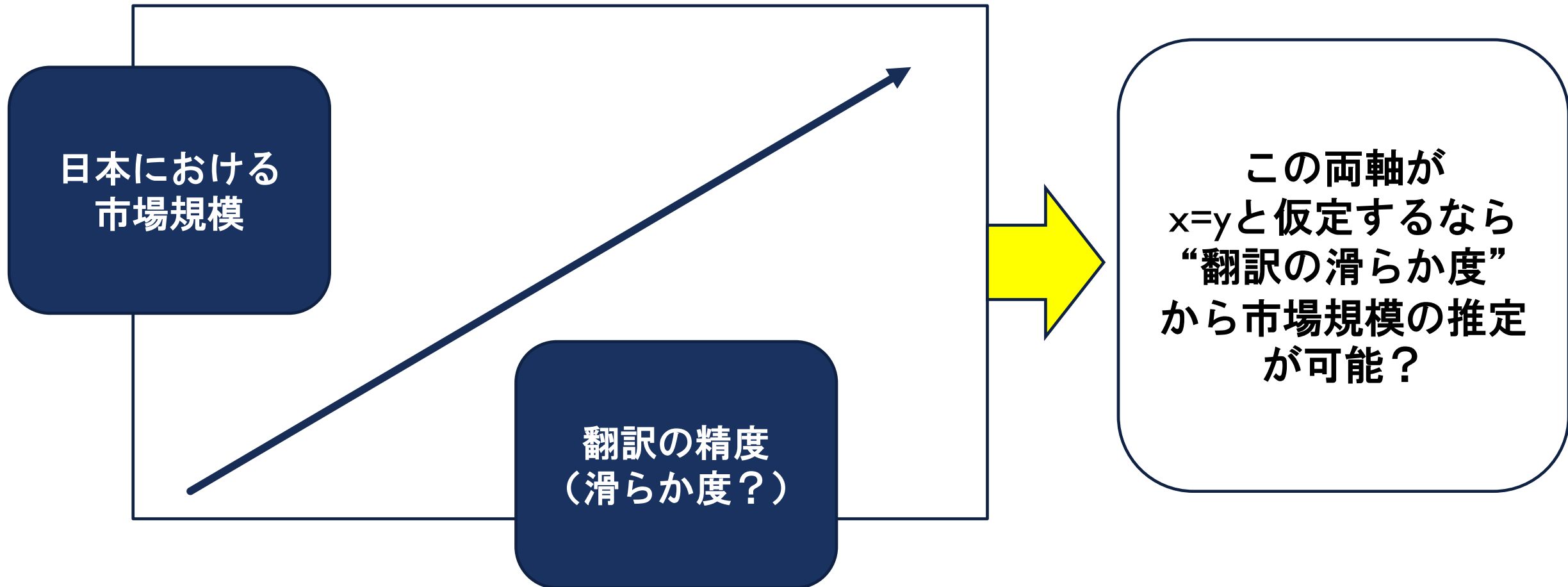


日本における市場規模が不明な場合に、
翻訳エンジンの精度から測ることができるか？



どうやって？

推定方法



推定方法

- y軸にとった **“翻訳の滑らか度”** とは？
- 誤訳を数値化する一般的な方法とは異なる。
(誤訳の頻度や重大度)
- 専門家が見て **“しっくりくる”** 文章になっているかどうかなので、**主観に左右される。**

推定方法

- ①ポストエディットの作業時間の比較→減少
- ②日本語訳のGoogle検索ヒット数→増加
- ③用語集のGoogle検索ヒット数→増加

 比較が必要なので、ある時点のみで測定するのは困難

結論

市場への浸透とともに翻訳の精度が上がる



日本における市場規模については、訳者の感覚で捉えることはできるが、数値については今後も検討の余地がある