

第4回 日本法令の国際発信の推進に向けた官民戦略会議

機械翻訳（AI翻訳）について

令和5年 3月 6日

凸版印刷株式会社

INDEX

1. 調査研究で実施した事項
2. 翻訳システム
3. 課題への対応方法
4. 翻訳結果
5. 新たな課題
6. 今後の法令外国語訳についての提案
7. デモンストレーション

前段：令和3年度に抽出された課題

課題 1 不適切な主語の補い

(例) 裁判外紛争解決手続の利用の促進に関する法律

【日本語原文（第二十三条）】

四項 法務大臣は、第一項又は第二項の規定により認証を取り消したときは、その旨を官報で公示しなければならない。

【機械翻訳による翻訳結果】

(4) If the Minister of Justice rescinds certification pursuant to the provisions of paragraph (1) or (2), the Prime Minister must issue public notice of this in the official gazette.

主語のない日本語から英語への機械翻訳過程で、「首相」を意味する「Prime Minister」が主語として補完されているが、正しくは、「法務大臣」を指す語句（単数の”they”や”That Minister”等）にすべきであり、主語が誤って補われている。

課題2

訳語の不統一、「法令用語日英標準対訳辞書」・「法令翻訳の手引き」への非準拠

(例) 債権管理回収業に関する特別措置法施行規則

【日本語原文】

第七条 法第十条第一項に規定する届出をする者は、別紙様式第十号により作成した廃業等届出書に、次に掲げる書面を添付して、法務大臣に提出しなければならない。

- 一 届出をする者が法第十条第一項各号の区分に応じその号に定める者である旨を証する書類
- 二 届出に係る事項を記載した登記事項証明書又はこれに代わる書面

【機械翻訳による翻訳結果】

Article 7 (1) A person who makes a notification as prescribed in Article 10, paragraph (1) of the Act must submit to the Minister of Justice a written notification of discontinuance of business, etc. prepared using appended Form 10, attaching the following documents :

- (i) a document certifying that the person who makes the notification is a person prescribed in each item of Article 10, paragraph (1) of the Act according to the category set forth in the item ;
- (ii) A certificate of registered matters containing the matters pertaining to the notification or any other document in lieu thereof ;

- ・各号の先頭文字は小文字とする。
- ・各号の末尾は、最後の号以外は「;」，最後の号は「.」をつける。
- ・最後からひとつ前の号の末尾は「; and」または「; or」
- ・号の表記は小文字のローマ数字、 といった翻訳ルールと整合が取れていない。

1. 調査研究で実施した事項

翻訳システムの構築

前記2つの課題を重点解決事項として、翻訳エンジンに翻訳品質改善を実現するプログラムを実装し、「翻訳システム」を構築



翻訳サンプルの生成

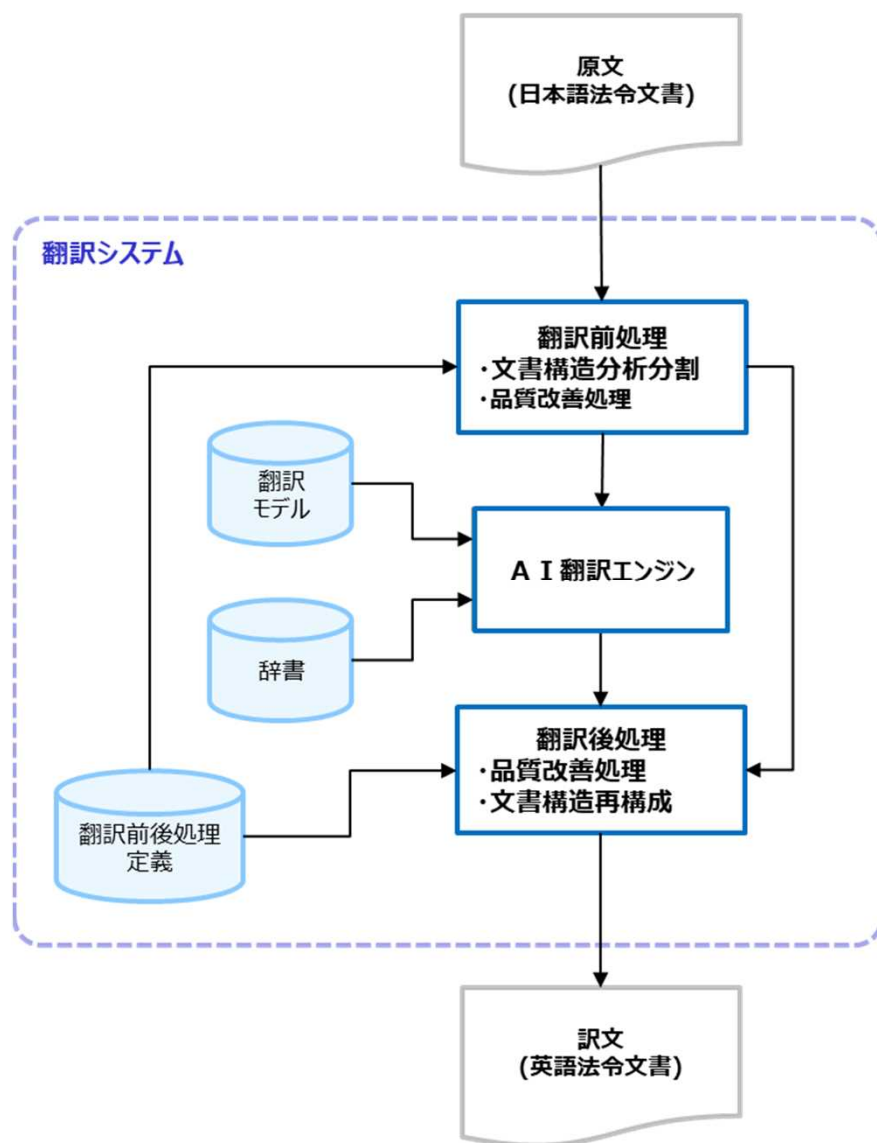
法令データを、構築した翻訳システムで実際に機械翻訳し、翻訳サンプルを生成



翻訳結果の確認・分析

翻訳サンプルを用いて、翻訳結果の考察及び、より効率的かつ効果的なAI学習の強化の方法や、より精度の高い翻訳品質改善のためのプログラムについて調査・検討を実施

2. 翻訳システム



本調査研究では、令和3年度に見られた課題を解決するための仕組みとして、新たな**翻訳システム**を開発し、調査を行った。

翻訳システムは翻訳原文に対し、翻訳結果が出力される一連の仕組みを指す。

図は今回構築した翻訳システムの概要である。

翻訳エンジンとは

翻訳エンジンは、深層学習（ディープラーニング）を利用したニューラル機械翻訳（＝AI翻訳）を用いた。
深層学習とは、脳の神経回路の一部を模したニューラルネットワークを用いた機械学習の手法である。

翻訳モデルとは

AI翻訳では、対訳文を使用して、深層学習を行うことによって、翻訳モデルと呼ばれるデータベースを構築する。
構築された翻訳モデルを利用して翻訳エンジンが翻訳処理を実行する。
本調査研究では、自動翻訳技術の研究開発を推進している「国立研究開発法人情報通信研究機構（NICT）」が構築した「法令契約モデル」を主に活用した。

3. 課題への対応方法

①追加学習による改善

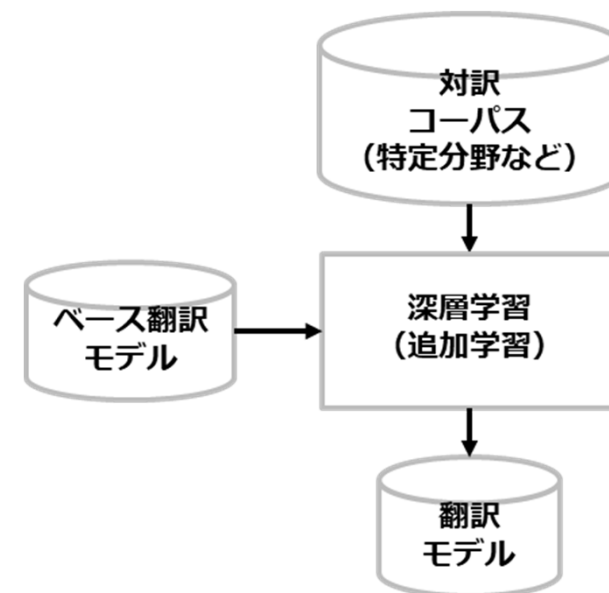
あらかじめ構築済みの翻訳モデルをベースに、特定の対訳文を追加学習（アダプテーション）することで、特定分野に一層適応した翻訳モデルを構築する手法を用いた改善策

②前後処理プログラムによる改善

翻訳品質の改善を目的に、翻訳エンジンの前段階と後段階に組み込む機械処理を用いた改善策



本実証研究では、この2つの改善方式が含まれる「**翻訳品質改善プログラム**」を実装した。



課題 1 不適切な主語の補いの解決

課題の一例：日本語側における「大臣」と英語側における“Minister”の出現数が一致しない

日本語	英語
4 内閣総理大臣は、政令で定めるところにより、定期的に、及び任命権者の求めがある場合その他必要があると認める場合には随時、適格性審査を行い、幹部候補者名簿を更新するものとする。	(4) The Prime Minister is to implement the eligibility screening and renew the executive candidates list periodically, whenever an appointer requests to implement it and whenever the Prime Minister finds it to be necessary, pursuant to the provisions of Cabinet Order.

解決策

本業務では、JLTから収集した最新の対訳文から、原因となる対訳を含む対訳文を取り除き、対訳データの質を向上させた上で、前述のNICT法令契約モデルに追加学習を実施した。



機械翻訳の普遍的な課題であるため、追加学習の効果により、不適切な補いの発生を根本から防ぐ。

課題2 訳語の不統一、「法令用語日英標準対訳辞書」・「法令翻訳の手引き」への非準拠の解決

課題の一例：条・項・号の英訳表記

- ・項の訳抜け＝（１）表記が抜ける
- ・号の表示については小文字で記載すべきであるが大文字になる
- ・1項しかない場合、（１）の表記は不要であるところ、（１）と入る。

解決策①

日本語側から条・項・号を取り除き、本文のみにした上で、**追加学習を実施**

解決策②

上記の本文のみと条・項・号記述ルールに従って取り除いた条・項・号を別途翻訳し、**本文翻訳結果と条・項・号の翻訳結果を統合するプログラムを実装**



追加学習と前後処理プログラムを組み合わせることで、法令翻訳に特有の翻訳課題の解決を可能とする。

4. 翻訳結果

課題 1 不適切な主語の補いの解決

結果

不適切な主語の補いは発生せず、改善されていた。

改善事例 相続等により取得した土地所有権の国庫への帰属に関する法律第 13 条第 4 項

日本語原文	翻訳品質改善プログラム 実施前サンプル	翻訳品質改善プログラム 実施後サンプル
<p>4 法務大臣は、第一項の規定により第五条第一項の承認を取り消したときは、法務省令で定めるところにより、その旨を同項の承認を受けた者に通知するものとする。</p>	<p>(4) If the Minister of Justice rescinds the approval referred to in Article 5, paragraph (1) pursuant to the provisions of paragraph (1), <u>the Prime Minister</u> is to notify the person that has obtained the approval referred to in that paragraph of this pursuant to the provisions of Ministry of Justice Order.</p>	<p>(4) Having rescinded the approval referred to in Article 5, paragraph (1) pursuant to the provisions of paragraph (1), <u>the Minister of Justice</u> is to notify the person that has obtained the approval referred to in that paragraph of this, pursuant to the provisions of the Order of the Ministry of Justice.</p>

課題2

訳語の不統一、「法令用語日英標準対訳辞書」・「法令翻訳の手引き」等への非準拠の解決

課題例：条・項・号の英訳表記

結果

概ね改善されていた。一部未改善項目に関しても解決策が明確であった。

改善事例 債権管理回収業に関する特別措置法施行規則 第6条第1項第1号（号の表記は小文字のローマ数字）

日本語原文	翻訳品質改善プログラム 実施前サンプル	翻訳品質改善プログラム 実施後サンプル
二 合併の経緯等を説明した書面	(I) a document explaining the particulars, etc. of the merger ;	(i) a document explaining the particulars, etc. of the merger ;

5. 新たな課題

①対訳文の除去について

- ・ 正確性の高まった最新のJLT対訳から、原因となる対訳文の除去を実施した上で追加学習を行ったことにより、課題が改善した。一方、追加学習による全体の訳質変化や一部悪化（対訳除去の影響の可能性）も認められた。

このことから、追加学習、とりわけ新たな対訳を用いて行う追加学習に関しては、実施に向けて慎重な判断が必要であり、実施の際は入念な調整が必要である。

②ルール同士の抵触について

複数の改善処理が抵触し合った場合、翻訳結果に悪影響を与えることが確認された。対策方針の検討には更なる調査を要するが、取り扱い方法のルールを整理して改善プログラムに採用すること等により改善が期待できる。



翻訳システムの導入や追加学習、新ルール導入にあたっては、翻訳品質についてのテストを複数の法令を用いて実施し、入念な調整が必要である。

6. 今後の法令外国語訳についての提案

新たな課題も発見されたが、翻訳の品質は、翻訳品質改善プログラムを搭載することにより大幅に向上しており、法令外国語訳業務において、英訳法令の原案作成にAI翻訳が有用であることが確認できた。

ただし、日本語としても構造が分かりにくい文については、英文として成立していたとしても、係り方（解釈）が正しいか等について人の目で確認する必要があるほか、大文字/小文字、単数型/複数型、文脈に応じた訳語の選定等については、機械的に判断することは難しいと考えられる。

このことから、人による確認・修正を前提とし、チェックの在り方についても、充実させることを検討すべきである。



翻訳品質改善につながった翻訳改善プログラムとその継続的な調整に加え、法令外国語訳における翻訳品質のチェックの在り方を充実させた上で、英訳法令の原案の作成に臨むこととすれば、英訳法令の原案作成までの期間を抜本的に改善でき、英訳法令の原案作成期間短縮の結果及びチェック体制の充実等によって、公開までの期間も短縮できると考えられる。

7. デモンストレーション

TOPPAN