

自動車法規文の自動翻訳をニューラル技術で高精度化 ～トヨタとの共同研究を通じ、英日・中日翻訳の実用度が向上～

【ポイント】

- 自動車業界からトヨタが翻訳バンクに協力、翻訳データを提供
- 自動車法規を対象とした翻訳をニューラル英日翻訳システムで24%実用度向上
- 数万点に及ぶ部品のメーカーを含めた裾野の広い自動車業界全体への波及効果は絶大

国立研究開発法人情報通信研究機構(NICT、理事長: 徳田 英幸)は、トヨタ自動車株式会社(トヨタ、代表取締役社長: 豊田 章男)と自動車法規を対象としたニューラル英日・中日自動翻訳の共同研究を実施し、実用度向上を確認しました。これにより、自動車業界全体でのニューラル自動翻訳の活用が大きく期待できます。

今後は、本技術を法規文以外のマニュアル等の多種多様な文献に展開するとともに、多言語化を進め、自動車産業で必要となるあらゆる翻訳を高効率化することを目指します。

【背景】

自動車及び自動車部品の設計・製造・輸出に当たっては、安全確保・環境保護等の様々な観点から、各国が定める法規の要件を満足することが求められています。そのためには、頻繁に改訂される各国の法規情報を迅速かつ正確に理解する必要があり、現状は、人による翻訳のため、精度確保には時間が掛かっています。情報の高速・高精度な翻訳は、トヨタをはじめとする自動車業界の業務効率化の鍵となっています。

NICTでは、2017年から総務省と連携して、ニューラル翻訳¹技術に不可欠な翻訳データを集積する翻訳バンク²を運営し、製薬会社をはじめとして多数の組織からデータの提供を受けて、翻訳データの集積・活用を進めてきました。また、トヨタとは2018年6月から「自動車法規文章の自動翻訳エンジンの研究開発」で共同研究を進めてきました。

【今回の成果】

NICTは、トヨタが提供した自動車法規に関する英日・中日翻訳データを基に、汎用英日・中日ニューラルネットワーク翻訳エンジンのアダプテーション³を行い、実用度を向上させました。

トヨタが自動車法規についてアダプテーション後の実用度を評価したところ、英日翻訳について、アダプテーション前と比べ、自動翻訳エンジンの約24%、中日翻訳については、約11%の実用度向上が得られました。これを受けて、トヨタでは、実用度向上の検討を継続することとなりました。

各国語の法規の翻訳は、自動車にとどまるものでなく輸出される全ての生産物に対して存在し、その翻訳の高速化・高精度化は同様に不可欠となり、翻訳バンクによって実現されていくことが期待されます。

【今後の展望】

今回は、英日・中日翻訳の片方向のみでしたが、今後は、①双方向に、更には、多言語にすること、②オーナーズマニュアル等の多種多様な文献に展開すること、③自動車産業全体に広げること、④広域にわたる輸出産業に広げることを目指します。

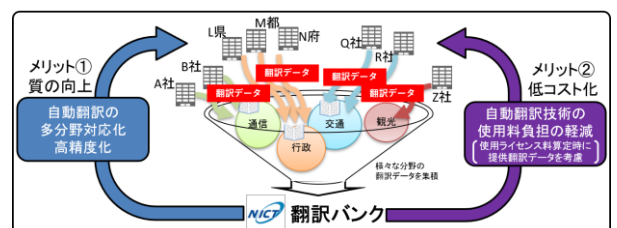


図1 翻訳バンクスキーム

<原文と自動翻訳結果のサンプル>

法規原文	人による翻訳	ニューラル翻訳
This Australian Design Rule prescribes requirements for the number and mode of installation of lighting and light signalling devices on motor vehicle other than L-group vehicles.	本オーストラリア設計規則は L グループ車両以外の自動車への灯火および灯火信号装置の数と取り付け方法に関する要件を規定する。	本オーストラリア設計規則は、L グループ車両以外の自動車への灯火および灯火信号装置の数および取り付け方法に関する要件を規定する。
本标准规定了电动汽车传导式车载充电机(以下简称车载充电机)的术语与定义、参数、要求、试验方法、检验规则及标示、包装、运输和储存。本标准适用于纯电动汽车和可外接充电的混合动力电动汽车用的车载充电机	本標準では電気自動車の伝導式車載充電器(以下、車載充電器という)の用語と定義、パラメータ、要求、試験方法、検査規則及びマーク、梱包、輸送及び保管について規定している。本標準は、純電動自動車及び外部接続充電式ハイブリッド電気自動車用の車載充電器に適用する。	本標準は電気自動車伝導式車載充電器(以下、車載充電器と略称する)の専門用語と定義、パラメーター、要求、試験方法、検査規則及び表示、包装、輸送及び保存を規定した。本標準は純電動自動車と外部充電が可能なハイブリッド電動自動車用の車載充電式充電器に適用される。

<用語解説>

*1 ニューラル翻訳(NMT)

脳の神経回路を模したニューラルネットワークを用いた自動翻訳技術。膨大な翻訳データを用いてトレーニングしたニューラルネットワークで翻訳することで、従来の翻訳技術よりも高い翻訳精度が確認されている。

<https://www.nict.go.jp/press/2017/06/28-1.html> (NICT)

*2 翻訳バンク

ニューラル技術による自動翻訳の精度向上には、アルゴリズムの改良に加えて、翻訳データの質と量の影響も大きく、高品質翻訳データの大量の確保が重要となる。NICT は、総務省と共に翻訳データを集積する「翻訳バンク」を運用し、日本語の翻訳技術の多分野化・高精度化に取り組んでいる。

<http://h-bank.nict.go.jp/index.html>

<https://www.nict.go.jp/press/2017/09/08-1.html> (NICT)

提供された翻訳データは、既存の翻訳データと併せて、ニューラルネットのトレーニングに活用され、技術移転される。最新版の翻訳精度は、NICT の開発した音声翻訳アプリ VoiceTra⁴ や文字ベースの自動翻訳システム TexTra⁵ で自由に確認できる。

*3 アダプテーション

翻訳バンクで追加した翻訳データを用いて、精度を改善するようにトレーニング済みの NMT のニューラルネットワークを更に調整すること。

*4 VoiceTra

言葉の壁で困らない社会実現を目指して NICT が開発した多言語音声翻訳アプリ。NMT を用いて、観光地での旅行会話はもとより、病院、商業施設といった様々なシーンで精度の高い音声翻訳を提供している。世界 31 の言語に対応。

<https://voicetra.nict.go.jp/>

*5 TexTra

NICT では、文字入力用の NMT を TexTra と名付けて公開している。公開サイト「みんなの自動翻訳@TexTra」では、コピー・ペーストしたり、サイト上の翻訳エディタを利用したり、ワードやパワーポイントのファイルを直接翻訳したり、API (Application Programming Interface) を介してプログラムから利用するなど、様々な方法で翻訳精度を試すことができる。

<https://mt-auto-minhon-mlt.ucri.jgn-x.jp/> (NICT)

※ 翻訳バンク、VoiceTra、TexTra は NICT の登録商標です。

< 本件に関する問い合わせ先 >

国立研究開発法人情報通信研究機構
先進的音声翻訳研究開発推進センター
知能科学融合研究開発推進センター(兼務)
隅田 英一郎
Tel: 0774-98-6350
E-mail: textra-info@khn.nict.go.jp

< 広報 >

広報部 報道室
廣田 幸子
Tel: 042-327-6923
Fax: 042-327-7587
E-mail: publicity@nict.go.jp