

AI技術に関するオープンイノベーション型研究プロジェクトの推進

■概要

連携推進室では、AI技術に関するオープンイノベーション型研究プロジェクトとして、先進的音声翻訳研究開発推進センターと共に翻訳バンクを運用している。平成30年度における翻訳バンクの運用に関する主なトピックスは、1.翻訳データの集積、2.「第2回自動翻訳シンポジウム～自動翻訳と翻訳バンク～」の開催、3.翻訳データのアダプテーション用サーバの運用開始である。

■平成30年度の成果

1. 翻訳バンクの概要及び翻訳データの集積（図1）

NICTでは、2020年までに多言語音声翻訳技術の社会実装を目指す「グローバルコミュニケーション計画」の下、AI技術で多用される深い階層構造を持つニューラル

ネットワークを用いた自動翻訳技術（ニューラル翻訳）の研究開発に取り組んでいる。ニューラル翻訳による自動翻訳の精度向上のためには、ニューラルネットワークのアルゴリズムの改良に加えて、様々な分野の翻訳データを大量に確保することが重要である。一方、国の機関や都道府県、市町村等の地方自治体、民間企業には、これまで多言語で作成された書類、観光案内等のパンフレット、業務説明資料、取扱説明書等の様々な文書が多く存在している。情報通信審議会情報通信技術分科会技術戦略委員会第3次中間答申（平成29年7月20日）においては、それら多様な文書から同じ意味を持つ単語、文の「対」を取り出し、仮にNICTにこれらのデータを集約できれば、これらのデータを組み込むことが可能となり、翻訳システムの精度向上に資する旨が述べられている。これを踏まえ、NICTは総務省と連携し、平成29年9月、オールジャパン体制で翻訳データを集積する「翻訳バンク」の運用を開始している。今後、翻訳バンクによりNICTに集積した翻訳データを活用することにより、我が国発の翻訳技術の多分野化・高精度化が進展することが期待されている（図2）。

平成30年度においては、府省庁からの翻訳データの集積をはじめ、翻訳バンクはオールジャパン体制による取組へと進んでいる。様々な分野で高精度翻訳（図3）を実現することで「言葉の壁」をなくし、日本を『世界で最も多言語コミュニケーションが容易な国』にするこ



図1 翻訳バンクのロゴマーク

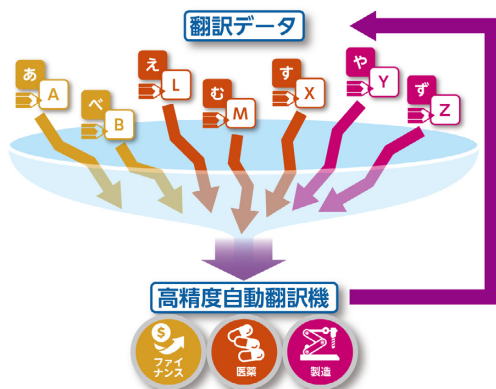


図2 翻訳バンクのコンセプト



図3 高精度翻訳の分野

とによって、日本の経済・社会の活性化に貢献することを目指している。

2. 「第2回自動翻訳シンポジウム～自動翻訳と翻訳バンク～」の開催(図4)

翻訳バンクの認知度向上を図るため、前年度に続き、NICTと総務省が主催する「第2回自動翻訳シンポジウム～自動翻訳と翻訳バンク～」を内閣府、法務省、外務省、特許庁の後援の下で平成31年3月6日に開催した。

冒頭の総務大臣政務官の主催者挨拶に続いて、AI、自



図4 「第2回自動翻訳シンポジウム～自動翻訳と翻訳バンク」会場内の模様

動翻訳及び関連法規に関する講演が実施された。また、「AI時代の研究開発、ビジネス及び法律問題」と題したパネルディスカッションも実施され、翻訳バンクがデータ収集の優れたモデルであることや、他の分野のデータについても大規模にデータを収集することの重要性等について熱い議論が行われた。一方、会場に併設した展示コーナーでは多言語音声翻訳に関係する民間企業による10件の展示が行われ、好評を博した。

参加者数が前回から100名増の300名を超える大盛況のうちに閉幕した。東京オリンピック・パラリンピック競技大会を翌年に控え、多言語音声翻訳技術への関心は一層高まりつつあることを実感する有意義なシンポジウムとなった。

3. 翻訳データのアダプテーション用サーバの運用開始

翻訳データをNICTに提供するメリットをより明確化するため、NICTの自動翻訳技術の利用に際して多数の翻訳データを提供した場合には、提供者自らが翻訳エンジンに翻訳データをアダプテーションすることができるというスキームを開始した。これにより、各提供者は自らが提供した最新の翻訳データにおける翻訳傾向を踏まえた翻訳エンジンを利用できるようになった。