



平成 29 年 9 月 8 日
 総 務 省
 国立研究開発法人情報通信研究機構

『翻訳バンク』の運用開始

ー自動翻訳システムのさらなる高精度化に向けて、様々な分野の翻訳データを集積ー

総務省と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、自動翻訳システムの様々な分野への対応や高精度化を進めるため、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』の運用を開始します。これにより、社会・経済活動のグローバル化が進む中で我が国の国際競争力の強化に貢献します。

【ポイント】

- オール・ジャパンで様々な分野の翻訳データを集積する『翻訳バンク』の運用を開始
- ニューラル機械翻訳での活用等に向けて翻訳データを集積し、多分野対応や高精度化を実現
- 翻訳データを提供して頂く方にもメリットのある仕組みを導入

1 背景・課題

総務省と NICT は、世界の「言葉の壁」をなくすことを目指すグローバルコミュニケーション計画^{*1}を推進しており、その一環として NICT は音声翻訳（“VoiceTra^{*2}”と呼ぶ）とテキスト翻訳（“TexTra^{*3}”と呼ぶ）の研究・開発・社会実装を進めています。これまで NICT においては翻訳の高精度化に必要な翻訳データの集積（補足説明の表 1）に取り組むとともに、2017 年 6 月からニューラル機械翻訳技術^{*4}の導入等を進めておりますが、翻訳技術を活用する分野によっては翻訳データが足りないことが課題となっていました。

2 翻訳バンクの運用開始

そこで、総務省と NICT は、様々な分野における翻訳データの集積に向けて、NICT が様々な分野の翻訳データを集積して活用する『翻訳バンク』の運用を開始することにしました（図 1）。

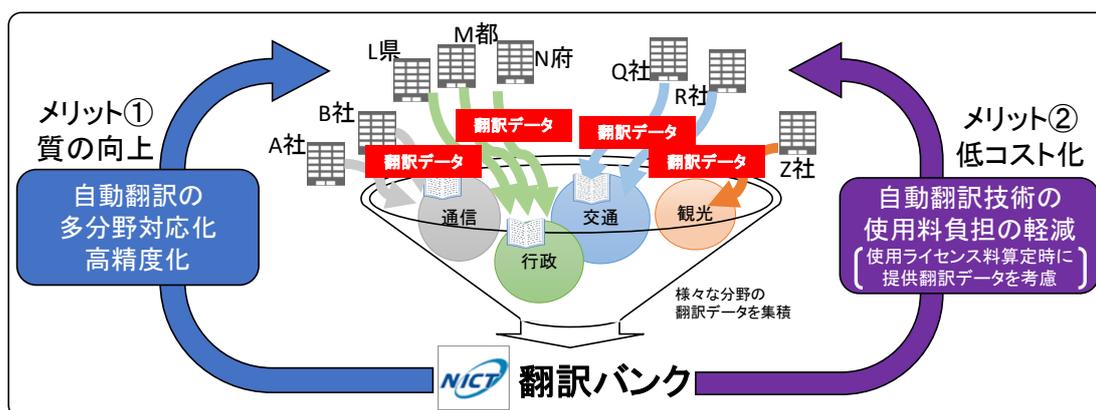


図 1 『翻訳バンク』のコンセプト

『翻訳バンク』の開始にあたり、データを提供して頂く方のメリットを明確化するため、NICT の自動翻訳技術の使用ライセンス料の算定の際に、提供が見込まれる翻訳データを勘案して負担を軽減する仕組みを用意しました。

3 今後の展望

総務省と NICT は、様々な分野で高精度翻訳を実現することで「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』にすることによって、日本の経済・社会の活性化に貢献します。

<用語解説>

*1 グローバルコミュニケーション計画(2014年4月)

世界の「言葉の壁」をなくしグローバルで自由な交流を実現するために、音声翻訳技術の多言語化・多分野化を推進し、2020年までに社会実装を目指すと発表。

http://www.soumu.go.jp/main_content/000285578.pdf(総務省)

*2 VoiceTra

NICTの開発した音声翻訳アプリ。31カ国語に対応。

<http://voicetra.nict.go.jp/>(NICT)

*3 TexTra

NICTの開発した最先端の文字ベースの自動翻訳システム。

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

*4 ニューラル機械翻訳技術

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。膨大な翻訳データから学習したニューラルネットワークを用いて翻訳することで、従来技術よりも高い翻訳精度が実現可能になることが確認されている。

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

<補足説明>

- 自動翻訳技術の性能向上のためには、NICTが研究開発において取り組んでいる翻訳アルゴリズムの改良に加え、翻訳データの質と量の確保が重要となっています。翻訳バンクでは翻訳データを集積して自動翻訳技術に活用することで、自動翻訳技術で対応できる分野を広げるとともに、さらなる高精度化を実現します。
- NICTの自動翻訳技術の使用ライセンス料の算定の際に、提供が見込まれる翻訳データを勘案して負担を軽減する仕組みを導入することで、翻訳バンクにおける好循環モデルを構築し、皆で自動翻訳技術を育てながら、高性能な自動翻訳技術を活用していくサイクルを確立します。
- 翻訳バンクの当面の目標として、100万文×100社=1億文の翻訳データの集積を目指します。
- 翻訳バンク Web サイト <http://www2.nict.go.jp/ais/h-bank.html>

表 1 収集する翻訳データのサンプル例

日本語	そのメインストリートといえるのが、祇園花見小路通である。
英語	The main street of Gion is Hanamikoji Street.
韓国語	그 메인스트리트라고 할 수 있는 것이 기온하나미코지도오리다.
簡体字	可以称之为主要街道的就是祇园赏花小路通。

(データ提供: ジャパンホッパーズ様)

- 現在 29 組織からデータ提供を受けています(2017年8月末現在)。
<http://voicetra.nict.go.jp/news/>

(連絡先)

【グローバルコミュニケーション計画に関すること】

総務省国際戦略局技術政策課研究推進室

担当 : 増子課長補佐、福川係長

電話 : (直通) 03-5253-5730

FAX : 03-5253-5732

Eメール : gcp.mic@soumu.go.jp

(連絡先)

【翻訳バンク・自動翻訳技術に関すること】

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

担当 : 隅田 英一郎

電話 : (直通) 0774-98-6350

Eメール : ltg-info@khn.nict.go.jp

【NICT 広報】

国立研究開発法人情報通信研究機構

広報部 報道室

担当 : 廣田 幸子

電話 : (直通) 042-327-6923

FAX : 042-327-7587

Eメール : publicity@nict.go.jp