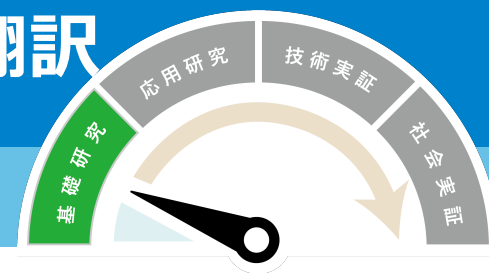


A08

# ユーザ生成テキストのニューラル機械翻訳



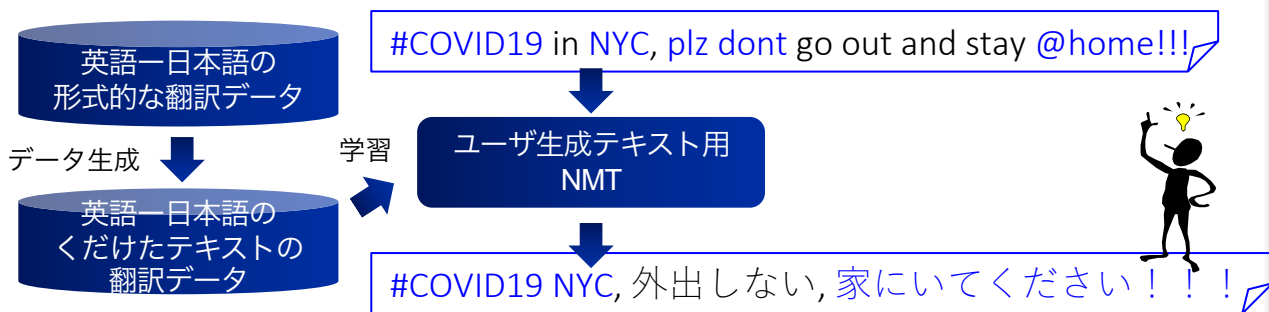
## 概要

既存の機械翻訳は、SNSなどに投稿されるようなテキスト（ユーザ生成テキスト）の翻訳が苦手です。我々は、翻訳データを自動生成して学習させることで、翻訳品質を向上させました。

## 機械翻訳の得手不得手



## 自動生成した翻訳データを用いたニューラル機械翻訳の学習



### 【お問合せ先】

国立研究開発法人情報通信研究機構 ユニバーサルコミュニケーション研究所  
 先進的音声翻訳研究開発推進センター 先進的翻訳技術研究室  
 bmarie@nict.go.jp, atsushi.fujita@nict.go.jp

## 特徴

- ユーザ生成テキストの翻訳データを人間が作成することなく品質を向上
- ユーザ生成テキストに見られる特徴を自動的に学習し、翻訳に反映（記号や略語の使い方）

## ユースケース

- SNSや個人ブログなどに投稿されるテキストの翻訳
- ネットショッピングサイトの商品レビューの翻訳

## 今後の展開

- 多様なユーザ生成テキストに対する性能の評価
- 様々な言語対への適用と評価
- 一定以上の性能が達成できたのちに技術移転・社会還元