

平成27年度 委託研究

課題 180

自治体向け音声翻訳システムに関する研究開発

研究計画書



1. 研究開発課題

『自治体向け音声翻訳システムに関する研究開発』

2. 研究開発の目的

自治体の様々な課題解決に向けてICTの利活用が期待されており、情報通信研究機構（以下、「機構」という。）においても、自治体のトップやICT業界の関係者らで構成する懇談会¹を立ち上げて、機構の研究成果を地域が抱える課題の解決に生かす方策を議論してきたところ、音声翻訳技術の自治体窓口等での利活用が期待されていることが確認された。

音声翻訳システムは、機構で研究開発した技術に基づき、旅行会話を対象として実用化²されている。しかしながら、旅行向けの音声翻訳システムを他分野あるいは他言語で利用しようとする、音声認識／翻訳できる語彙の不足など様々な問題が妨げとなって十分な性能が出ず、利活用可能なレベルには達していない。

本研究開発では既存の音声翻訳技術をベースとして、新たに実験・改良を繰り返して、自治体向けの音声翻訳システムの基盤技術を確立する。

この研究開発の成果物のうち、対訳辞書、対訳コーパス及び音声コーパスは、広く利活用可能な特質を持つことから、受託者が占有するのではなく、機構の帰属とし、機構を通して広く活用の道を開き、日本国民の利益に資することを目指す。コーパス構築に際しては、あらかじめ、プライバシー情報などが将来の利活用の妨げにならないような対策を立てる。

3. 採択件数、研究開発期間及び予算

採択件数：1件

研究開発期間：契約締結日から平成31年度までの5年間。

予算：各年度、総額150百万円(税込)を上限とする。

(提案の予算額の調整を行った上で採択する提案を決定する場合がある。)

4. 研究開発の到達目標

1) 自治体との強固な連携を確立する。

(1) 必須となる要件

研究成果は、広く自治体に利活用されることを目指すものであることから、音声翻訳に対するニーズのある自治体と連携し、必要なデータ収集への協力、実験への協力を確保し、最終的には、日本全国で通用するソリューションを追及する。ニーズはその自治体特有のものではなく、複数の自治体に共通して存在するものであること。

¹ <http://www.nict.go.jp/press/2014/01/30-1.html>

² <http://www.narita-airport.jp/jp/fun/app/naritra/>
http://pass.auone.jp/app/detail?app_id=5413400000002

2) 自治体窓口向け音声翻訳システムを実現する。

(1) 必須となる要件

- 旅行向けの音声翻訳をベースに自治体窓口等用の音声翻訳システムを構築する技術の研究開発
 - 音声認識、翻訳、音声合成等の機構の音声翻訳関連の要素技術との緊密な連携を前提とする³。
 - 日⇔3言語を対象とする自治体窓口業務などで利用可能な音声翻訳システムを構築する。3言語の内訳は、広く利活用されるように、受託者が自治体の現場のニーズを踏まえて提案することとする。
 - データ収集
 - ◇ 発音付対訳辞書（5万語以上）。品質管理のため、全件、合成音の聴取試験など性能評価を実施。
 - ◇ 対訳コーパス（20万文以上）。品質管理のため、翻訳モデル、言語モデルの学習、機構が指定するテストセット、翻訳系を使っての性能評価を実施。
 - ◇ 音声データ⁴（各言語4,000名以上、各話者0.1h以上）の収集・整備。品質管理のため、音響モデルの学習、機構が指定するテストセット、認識系を使っての性能評価を実施。
 - 上記と既存のデータの最適な混合手法
 - 実験とこれによる改良
 - ◇ 実験し、不足している語彙や表現を明らかにし、前記のデータを整備・拡張する。

(2) 実施することが望ましい要件

- (a) 1,741自治体（790市、23特別区、745町、183村）を対象にしたクラウドサービスのビジネスモデルの構築
- (b) 窓口業務を例にして、
 - (ア) 最適なユーザーインターフェースを構築するノウハウの確立
 - (イ) 必要となる用語や表現を効率的に収集する手法の確立

5. 研究開発の運営管理及び評価について

- 自治体向けの音声翻訳システムを念頭に、自治体で運用可能なシステムの基盤技術を確立すること。
 - 研究開発体制は、システムを研究開発する研究開発機関側と、それを利

³ NDA締結を経て、開示する。

⁴ 実施前に機構と調整する。

用する自治体側で構成される必要がある。

- 提案時に当該自治体を提示することが困難な場合は、研究開発開始までに、研究開発に参画する自治体を選定することも可能とする。この場合は、研究開発開始までに自治体が参画する旨を、提案書に記載ください。
- 研究開発に当たっては、機構の自主研究との連携を図ること。1 ヶ月に1回程度、機構の先進的音声翻訳研究開発推進センター Advanced Speech Translation Research and development promotion Center (ASTREC) と研究方針や進捗について定期的に打ち合わせを行うこと。
- 複数の機関が共同で受託する場合には、代表提案者が受託者間の連携等の管理運営を行うこと。
- 機構は研究開発の進捗状況等を把握するために、ヒアリングを実施することがある。また、平成29年度に中間評価(必要に応じて)を、平成31年度に終了評価を行う。終了評価の結果等を踏まえ、研究開発終了後に追跡評価を行う場合がある。