

グローバルコミュニケーションを実現させるコンテンツ処理基盤の研究開発

① 計画の概要

近年、「多言語音声処理技術」や「多言語翻訳技術」、「自然言語処理技術」の発展には著しいものがある。しかしながら、それらの技術を用いても人々の能力発揮や相互理解を阻害する「言葉の壁」や「知識の壁」は完全には解消されておらず、さらなる技術進화가求められている。例えばネット上のテキストを人々の知識源としてより効率的に活用するため、機械がそれらテキストの文脈を理解した上で行間を読むことを可能とする「自然言語処理技術」が期待されており、さらには、そうした技術によって、多数のテキストを横断的に読み、それぞれの意味内容を組み合わせることで知識を大量に獲得、蓄積することが期待されている。また、より人に近い、あるいは人を越えた音声インタフェースの実現のため、大騒音下での正確な音声認識や、人に比べて遜色のない音声合成を可能とする「多言語音声処理技術」も期待されている。さらに、こうした技術を組み合わせ、より高度なサービスの創出も期待されているところである。自然言語処理技術によって抽出した大量の知識や文脈情報と多言語音声処理技術を融合し、日常のトピックだけでなく、高度な専門的トピックについても雑談を行い、人の知識、知的能力を強化できる音声対話システムの実現が期待されている。また、物理的に高度な能力を持つロボットとの高度なコミュニケーションを可能とすることで、人の物理的能力の拡張が実現する。「多言語音声翻訳技術」とも融合することで、異なる言語の話者とそれぞれの言語で対話を行い、単なる翻訳を超え、文化、知識の差異も埋め、深い相互理解を実現する高度な異文化コミュニケーションツールの出現も期待されている。

本計画では、以上に述べたような技術を実現し、その社会実装を通して「言葉の壁」や「文化の壁」、「知識の壁」を越えて、人類の相互理解を促進し、人類の知識や能力を向上させることをめざす。

② 学術的な意義

学術的にも高度な以下の技術を開発して、社会実装を通じた形で国民に還元する。

(1) 自然言語処理技術

テキストに陽に書かれていない文脈情報である談話関係を抽出する談話関係抽出技術を開発し、これも利用しつつ高度かつ複雑な知識を大量のテキストから横断的に抽出可能とする。これらの技術により、対話システムにおいて、複雑かつ高度な質問への柔軟な回答を可能とし、さらには高度な知識を活用しつつ対話を行う柔軟な対話戦略や対話制御を実現する。最終的には、既存対話技術では不可能な、多様な話題に関する高度な雑談を可能とし、それらの雑談を通して人の知的能力、コミュニケーション能力を拡張できる対話システム等を実現する

(2) 多言語音声処理技術

実世界で利用される現場周辺の雑音混じった音声、いわゆる「現場音声」をスマホ等の幅広いデバイスで高精度に認識させる事を目的として、通常のマイクを用いる高精度な多言語現場音声認識技術を開発する。さらには、より自然な多言語音声合成技術も開発し、雑音が多い環境下でも人と違和感なくやり取りが可能な音声インタフェースを実現する。

(3) 多言語翻訳技術

高精度な翻訳技術を実現させるために文脈処理技術に加えて、1文単位ではない翻訳技術を開発する。また、より少ない言語資源で高精度な翻訳を行うための単言語コーパスや、コンパラブルコーパスによる高精度翻訳技術を開発する。この技術は既存の大規模コーパスが必要な翻訳技術からパラダイムシフトさせる次世代多言語翻訳技術を実現させることになる。さらに談話関係抽出技術と多言語翻訳技術を融合させて、高度な技術に関する翻訳や、翻訳における誤解を回避できる多言語翻訳技術によってグローバルな異分野・異文化コミュニケーションを実現する。

③ 国内外の動向と当該研究計画の位置づけ

大規模なテキスト情報から高度な意味的分析を行う技術は、国内において京都大学、東北大学等が類似の研究開発をしているが、NICTにおいて先進的かつ大規模なシステムが開発されている。AIスピーカー等の対話技術は、命令実行型の対話をルールベースで解釈するものがGoogle等で実用されているが、ルールベースの作り込みが不要な対話技術は、次世代の技術として基礎研究レベルである。

多言語音声処理・翻訳技術はGoogle等の海外のIT企業が先進的な研究開発を行っている。国内においてはグローバルコミュニケーション計画によって多言語音声翻訳技術の社会実装が加速している。しかしながら、現在の実用的な多言語音声翻訳技術の多くは1文単位での翻訳であり、会話全体の文脈を考慮した同時通訳技術等は基礎研究レベルである。

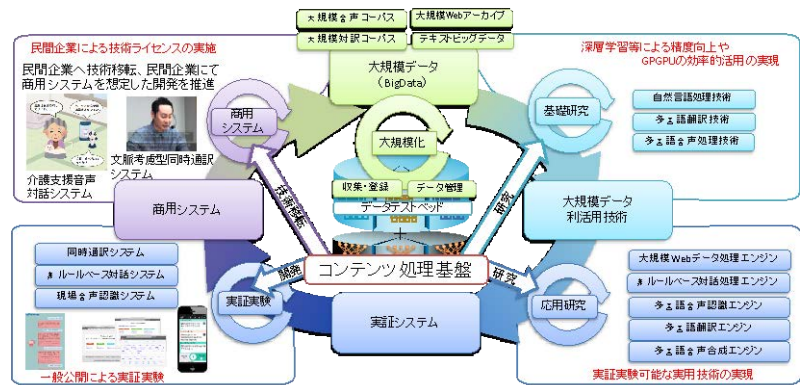


図1 グローバルコミュニケーションを実現するコンテンツ処理基盤の構築とコンテンツ処理基盤における研究開発

本計画は次世代の技術の研究開発をめざす。大規模テキスト情報等から抽出した知識を用いたルールを必要としない対話技術や文脈を考慮した同時通訳技術を2025年までに実現させると共に、その研究を支えるコンテンツ処理プラットフォームの構築を目標とするものであり、他に類を見ない先進的な研究計画である。

④ 実施機関と実施体制

本計画は、NICT のけいはんな拠点が中核機関として研究開発を推進する。「コンテンツ処理基盤の構築」については、NICT のけいはんな拠点に構築中のデータテストベッドのデータ処理機能を拡大することを目的として「コンテンツ処理基盤」を整備するものとする。また、「コンテンツ処理基盤における研究開発」については、NICT のけいはんな拠点の研究所・研究センターで研究開発を推進する。

2020年度まではNICTの既存の組織で対応することが認められているが、2021年度以降はNICTの次期中長期計画の中で推進することとなるため、今後の研究計画策定の中に本計画を実装していくこととなる。外部連携については、主にNICTの招へい研究者としてプロジェクトに参画している研究者を中心に推進する。既に東京大学、京都大学等の教授クラスの研究者がプロジェクトに参加しており、今後も本計画の研究開発に参画する予定である。

⑤ 所要経費

「コンテンツ処理基盤の構築」については、「コンテンツ処理基盤における研究開発」を推進するために、大型施設計画として計算機を整備する。このコンテンツ処理基盤を整備期間に合わせて6年間運用して「コンテンツ処理基盤における研究開発」を実施する。必要経費は以下の通りである。

【コンテンツ処理基盤の構築】 (2020年度～2025年度) 計1,100,000千円

大型施設整備：計算機 500,000千円 (2020年度のみ)

整備・運用にかかる人件費 技術者5名 60,000千円/年 計360,000千円 保守費 40,000千円/年 計240,000千円

【コンテンツ処理基盤における研究開発】計 (2021年度～2025年度)6,000,000千円

多言語対訳・音声コーパス作成費用 400,000千円/年 計2,000,000千円

Web アーカイブ・学習データ作成費用 200,000千円/年 計1,000,000千円

研究者人件費 50名 600,000千円/年 計3,000,000千円

⑥ 年次計画

2020年度に「コンテンツ処理基盤の構築」を行い、2021年度から「コンテンツ処理基盤における研究開発」を推進する。研究項目は「自然言語処理技術」「多言語音声処理技術」「多言語音声翻訳技術」の研究開発を推進する。

「コンテンツ処理基盤の構築」は、大型施設計画として2020年度にNICT内にGPGPUを持つ計算機を整備する。その後2025年度まで保守しながら運用を行い、「コンテンツ処理基盤における研究開発」として、「自然言語処理技術」「多言語音声処理技術」「多言語翻訳技術」の研究を推進する。

「自然言語処理技術」の研究開発においては、Web等の大量のデータから適切な応答を生成できる談話関係抽出技術を開発した上で、対話戦略、対話制御技術を開発する。これらの技術によってルールの記述が不要で大規模な知識源から自動的に適切な応答を生成することが可能となる。これにより、既存のルールベース対話技術の限界をブレイクスルーする対話システムを2025年度に実現させる。「多言語音声処理技術」の研究開発においては、実世界で利用される現場周辺の雑音が混じった音声、いわゆる「現場音声」を高精度に認識できる多言語現場音声認識技術を2025年度に実現させる。さらに人の音声と聞き分けが難しいレベルの自然な多言語音声合成技術を2025年度に実現させる。「多言語翻訳技術」の研究開発においては、文脈を考慮した長文翻訳技術と、それを応用した高精度な翻訳技術を2025年度に実現させる。また、開発した談話関係抽出技術や多言語現場音声認識技術、多言語音声合成技術、多言語翻訳技術を融合させることで、事件現場での記者の音声から講演音声まで実世界で利用できる同時通訳技術を2025年年度に実現させる。

⑦ 社会的価値

本計画を推進することで、既存の多言語音声翻訳技術では実現が困難な場面での同時通訳が実現されることになる。さらには大規模な知識に基づいて雑談も可能な音声対話技術が実現できる。これらの技術の応用範囲は極めて広く、我が国においてさらに多くの場面で言葉の壁がなくなる。さらに雑談も可能となる音声対話技術は介護等にも活用できるため、超高齢化社会を迎えつつある我が国においてICT・人工知能技術による介護支援が実現する。

本計画によって開発する技術は我が国固有の日本語の知的価値を向上させることは言うまでもない。また、開発した技術を企業の商用サービスに技術移転をすることで経済的・産業的価値も膨大なものとなる。さらには、言葉の壁をなくすことで、我が国で今後増えていく在留外国人の社会参画を容易にすることができ、安心・安全な社会を構築につながるため、本研究開発の効果については国民の理解も容易であり我が国におけるSDGsにも貢献できる。

我が国のAI戦略(有識者提案)においてAI中核センター群の一つとされているNICTを中心とする研究ネットワークによって研究開発を進めることは我が国のAI戦略にも適合している。

⑧ 本計画に関する連絡先

木俣 豊 (国立研究開発法人情報通信研究機構)