

## 1. 研究課題・受託者・研究開発期間・研究開発予算

- ◆研究開発課題名 自動翻訳の精度向上のための「マルチモーダル情報の外部制御可能なモデリング」の研究開発
- ◆副題 マルチモーダル情報理解と制御可能なテキスト生成の研究開発
- ◆受託者 (大)東京工業大学、(大)東京大学、(大)愛媛大学、東京都公立大学法人、日本放送協会、(株)時事通信社
- ◆研究開発期間 令和3年度～令和7年度(5年間)
- ◆研究開発予算(契約額) 令和3年度から令和5年度までの総額230百万円(令和4年度100百万円)

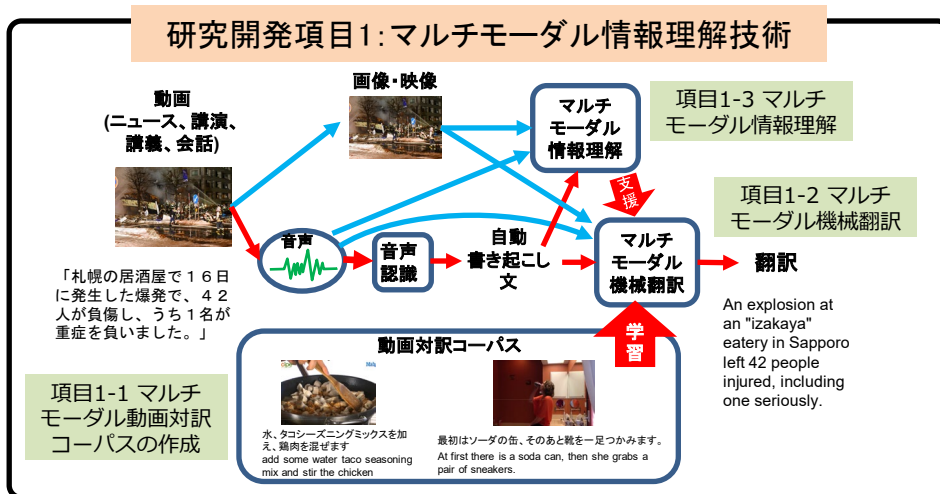
## 2. 研究開発の目標

人間同士および人間とコンピュータ間の円滑なコミュニケーションの実現に向けて、以下の研究開発を進める。

1. 文脈・状況を考慮しながら、言語、音声、画像、動画、表などによるマルチモーダルな情報を統合・解析する技術
2. 長さや焦点、スタイル、難易度、外部知識などで自然言語生成を柔軟に制御する技術

ニュース、講演、会議、対話などのドメインにおいて、翻訳や要約などのタスクに関連したコーパスを構築する。これらの成果を統合し、マルチモーダル情報を駆使しながら状況やニーズに応じて柔軟に通訳する自動通訳の実現を目指す。

## 3. 研究開発の成果

**項目1-1成果: マルチモーダル動画対訳コーパス**

- 既存の動画キャプションデータセットをベースとしてコーパスの作成を開始し、プレゼンテーション動画対訳コーパスでは約8時間・1500文対、インタラクション動画対訳コーパスでは約1時間・1300文対からなる部分データをそれぞれ作成した

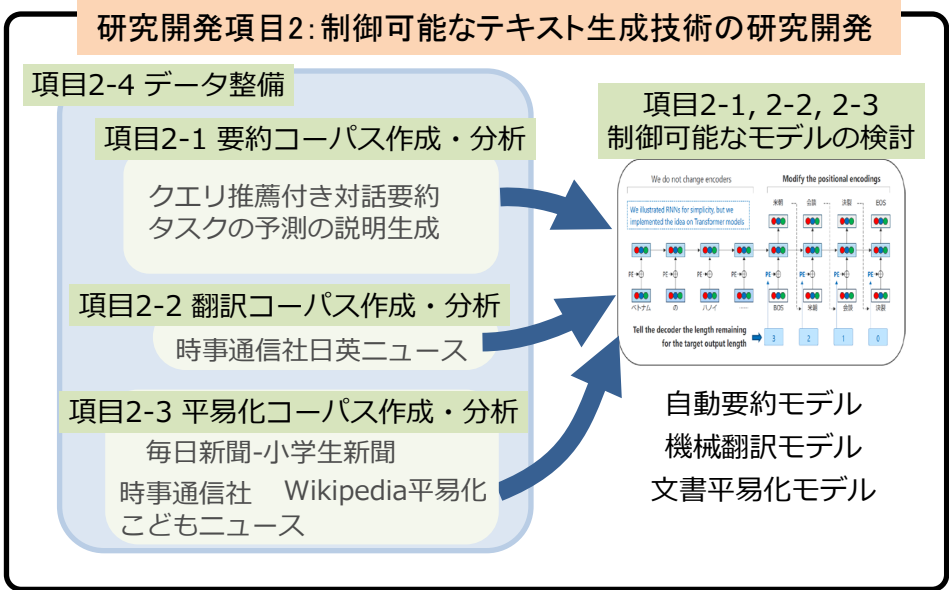
**項目1-2成果: マルチモーダル機械翻訳**

- IWSLT17対訳コーパスと原ビデオの対応付けを行い、音声付きマルチモーダル対訳コーパスの開発とマルチモーダル機械翻訳の性能評価を行った
- 拡散モデルを用い、入力画像を原言語文の説明に沿った画像に変換するマルチモーダル機械翻訳手法を開発した

**項目1-3成果: マルチモーダル情報理解**

- 学習に基づくデータ拡張や入力量子化により、テスト時の未知ノイズに対し頑健な画像認識手法を複数開発した
- カリキュラム学習により画像とテキストの接地を改善させる手法を開発した
- 外部知識グラフを用いて未来予測を行う画像キャプション手法を開発した

### 3. 研究開発の成果



- 項目2-1成果: 自動要約の制御に関する研究開発**
- クエリ推薦付き要約を提案し、その有効性を示した
  - タスクの予測の説明生成により、議論をしながら性能を向上させた
- 項目2-2成果: 翻訳の制御に関する研究開発**
- ニュースのライティングスタイルを考慮した日英ニュース翻訳用のテストデータを構築した
  - 機械翻訳の出力に対して編集量のパラメータを与えて生成結果を制御する手法を考案し、実験により有効性を確認した
- 項目2-3成果: スタイルの制御に関する研究開発**
- 統計的手法を用い、一般向け新聞記事と小学生向け新聞記事の自動対応付けを実施。日本語のテキスト平易化に関するコーパスを構築した
  - 改変・再配布等が可能なライセンス及びジャンルの観点からWikipediaを基にした平易化の人手タグ付けコーパスを構築した
- 項目2-4成果: データ整備に関する研究開発**
- 編集・配信DBより日英ペア記事、写真、子供ニュース等を抽出、配布した
  - 記事作成段階からデータの紐付けを行うフローで、記事セットを構築した

### 4. 特許出願、論文発表等、及びトピックス

国内出願	外国出願	研究論文	その他研究発表	標準化提案・採択	プレスリリース 報道	展示会	受賞・表彰
0 (0)	0 (0)	1 (1)	44 (43)	0 (0)	0 (0)	0 (0)	6 (6)

※成果数は累計件数、( )内は当該年度の件数です。

### 5. 今後の研究開発計画

マルチモーダル動画翻訳や自動要約、翻訳の制御、スタイルの制御などのコーパスの構築・拡張作業を進める。また、作成したデータについて順次公開の準備を行い、コンペティションのセットアップについても検討を進める。これらのコーパスに基づいて、マルチモーダル機械翻訳・情報理解、自動要約、機械翻訳、スタイル制御に関する新しい手法を開発する。