



2024年度 未踏IT人材発掘・育成事業 採択案件評価書

1. 担当PM

五十嵐 悠紀（お茶の水女子大学 理学部 情報科学科 准教授）

2. クリエータ氏名

青原 光（名古屋工業大学 大学院工学研究科 工学専攻 メディア情報プログラム）
伊藤 朝陽（名古屋工業大学 大学院工学研究科 工学専攻 ネットワークプログラム）

3. 委託金支払額

2,880,000円

4. テーマ名

スムーズな多言語交流を実現するためのARによる会話支援システム

5. 関連Webサイト

PredicTalk の App Store ダウンロード URL :

<https://apps.apple.com/jp/app/predictalk/id6476206495>

6. テーマ概要

本プロジェクトでは、AR グラスを活用してユーザの発話を支援しながらも、自然な会話の流れを維持できる iOS アプリケーション「PredicTalk」を開発した。PredicTalk は音声認識、次文予測、データ管理などの機能を持ち、Frame と Bluetooth で接続することでリアルタイムの音声データのやり取りを行う。単なる翻訳ツールではなく、「会話のスムーズな進行を支援するツール」としての役割を果たすものである。

7. 採択理由

本プロジェクトは、外国語で会話をする際に「対面で目を見て会話をしたい」という人間の本当にやりたいことを人間に残し、「次の言葉が出てこない」といった不足する語学力の問題を技術で支援するツールを開発するというものである。次に続く文章を言語モデルで予測し、ユーザが会話を継続できるようサポートする、という枠組みは、英会話だけでなく多言語への発展や、母語同士であっても会話のサポートを可能にするなど、様々な展開が期待できる

として採択した。新たなユーザ体験を提供するためのインターラクションデザインにも注力しながら取り組んで欲しいと期待する。

8. 開発目標

本プロジェクトではユーザが話している文章の続きの候補を AR グラスで視界内に表示することで、スムーズな多言語間の会話を支援するシステムを開発することを目標とした。開発システムでは、ユーザが会話をしている途中で単語や文法が思いつかず詰まってしまった際に、次に続く文章を言語モデルによって予測し、それをモバイルアプリケーションと連動した AR グラスに表示することで、ユーザが会話を継続できるようにするサポートをすることを目指した。

9. 進捗概要

本プロジェクトでは、AR グラスを活用することで、ユーザが会話をスムーズに継続できるよう支援するシステム「PredicTalk」を開発した。AR グラスには Brilliant Labs 製の Frame を使用することとした。Frame は、ハードウェア・ソフトウェアともにオープンソースであるため、開発者が独自の機能を柔軟に実装でき、軽量であり、透明なレンズであるため、自然な視界を保ちながら会話が可能である。また、Bluetooth によるスマートフォンとの無線接続も可能であり、本プロジェクトの目的に適したデバイスとして選定した。

PredicTalk では、ユーザが会話の途中で適切な単語や表現が思い浮かばず詰まってしまった際に、次の発話の候補を言語モデルがリアルタイムで予測し、AR グラスの視界内に候補を表示する（図 1）。これにより、ユーザは会話の相手の顔を見ながらスムーズに会話を継続でき、翻訳アプリケーションに依存することなく、自らの言葉で話すことが可能となる。PredicTalk は音声認識、次文予測、データ管理などの機能を持ち、Frame と Bluetooth で接続することでリアルタイムの音声データのやり取りを行う。次文予測には OpenAI の API を使用し、ユーザの発話を解析した上で、次に話すべき文章の候補を予測する仕組みを採用した。

本システムの構成は、iOS アプリケーションと Frame のデバイスドライバーの大きく 2 つの要素に分かれる。iOS アプリケーションは、直感的に操作できるデザインとなっており、初めて使うユーザでもスムーズに使うことができる。予測候補では複数の予測候補からユーザが選択できるように設計されている。3 つの短い文章でヒントが提示されるように設計したことで、会話中でも読みやすく、会話のヒントとなる単語が多く含まれるようになっている。また、ユーザが AR グラスをタップすることで新しいヒントを提示するようになっている（図 2）。

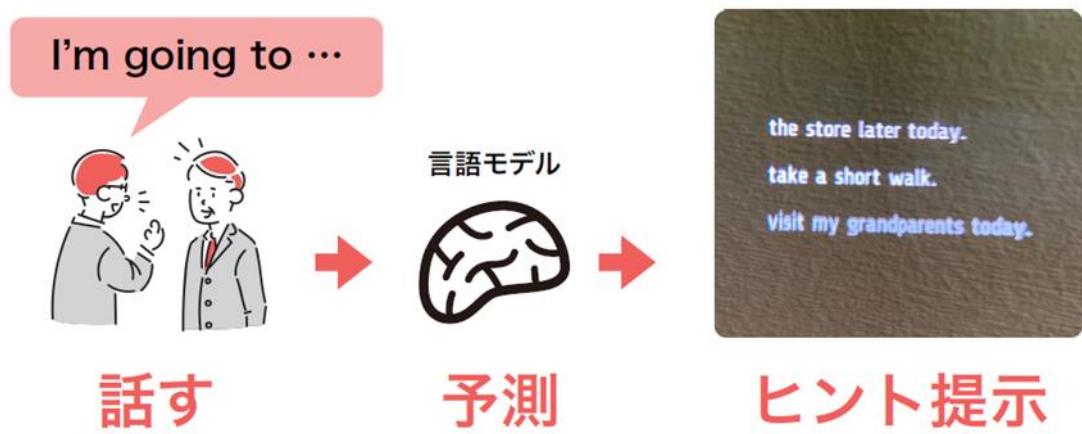


図 1：話していることを音声認識し、ユーザの発話の続きを予測し、AR グラス上に表示することで会話を支援する「PredicTalk」

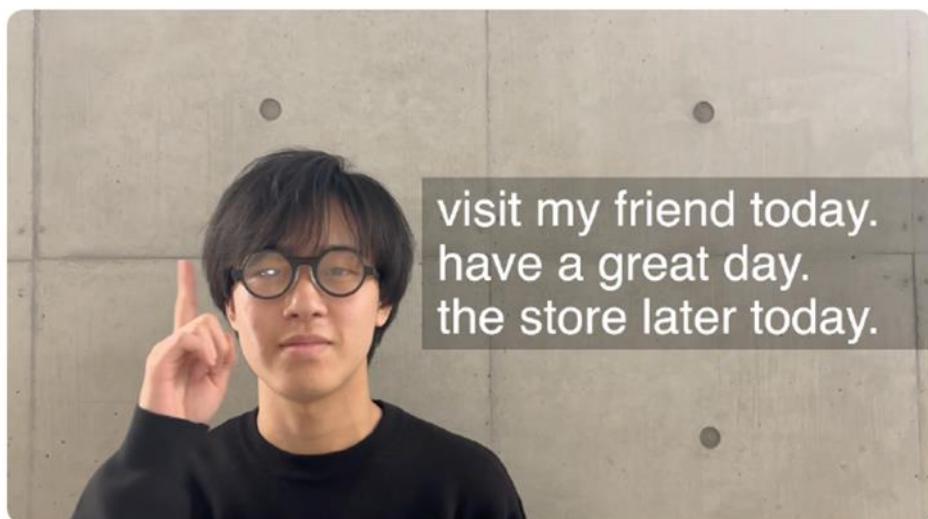


図 2：AR グラスをタップすることでヒントを変更できる

ユーザごとにカスタマイズすることもできる「設定画面」では、会話に詰まってから新しい予測を提示するまでの時間と、よりユーザに最適化された予測を実現するために予測モデルへ入力するユーザ情報を編集することができる（図 3）。これにより個人にあったテンポで会話を支援することができる。ユーザごとに経験や興味があるものといったユーザ情報を入力することで個人に適した予測も可能となった。



図 3：予測機能のカスタマイズ

プロジェクトを進めてきた中で大きな壁だったのは Frame の到着遅延である。予定されていた Frame がなかなか届かなかつたが、公開されている Frame の仕様やソースコードを読み解き、同じ基盤入手し、自作のシミュレータを開発するなど、試行錯誤を重ねながら進めてきた。また、Frame が手に入つてからは、実際に自分たちが使用する中で、「タイミングを調整できる方が良い」「3つの選択肢から選んだ方が会話しやすい」などの具体的な要件を検討し、仕様を定めていった。さらにユーザインタビューを実施し、ユーザ視点からのフィードバックを得ることで、より実用的な仕様へとブラッシュアップしていった。合宿を通じて多くの意見をもらい、必要な機能の優先順位を検討し、取捨選択を行ひながら、実装すべき機能を明確化した末出来上がつたのが PredicTalk である。

10. プロジェクト評価

本プロジェクトでは、「ユーザが話している文章の続きの候補を AR グラスで視界内に表示することで、スムーズな多言語間の会話を支援するシステムを開発する」という当初の目標を達成することができた。開発したシステムでは、ユーザが会話の途中で単語や文法が思い浮かばず詰まつてしまつた際に、言語モデルによって次に続く文章を予測し、それをモバイルアプリケーションと連動した AR グラスに表示することで、会話の継続を支援する仕組みを実現した。外国語でのコミュニケーションを技術で円滑にサポートする世界をみせた点を評価する。

11. 今後の課題

すでに形にはなつているがもっと一般に広く使ってもらうために是非頑張つて欲しい。方向性としては、英語での外国語コミュニケーション支援については

示すことができたが、もっと他の言語や「コミュニケーション支援」としての話題の提供など他の使い方の可能性も模索できると考える。ぜひ、今後も多言語化や一般公開を目指して頑張って欲しいと期待する。