

令和 6 年度研究開発成果概要書

採 択 番 号 06301

研究開発課題名 会話 AI エージェントとの高臨場感インタラクション体験実現のための XR 通信基盤の研究開発

(1) 研究開発の目的

本研究開発では、Beyond 5G で実現される「超高速・大容量」「超低遅延」「超多数同時接続」「自律性」「拡張性」「超安全・信頼性」「超低消費電力」通信インフラを基盤として、タブレット上の Web ブラウザや軽量の VR・AR ゴーグルなどあらゆるモバイルデバイスで手軽に体験できる会話 AI エージェントとの「breathhtaking な（息を呑むような）」高臨場感のある会話体験を実現することを目指す。本研究開発は以下の 3 つの研究開発項目から構成される。

- 研究開発項目 1：スケールアウト可能な会話 AI エージェントコンテナの開発：リッチクライアント型・シンクライアント型それぞれの環境でスケールアウト可能なアーキテクチャを設計・実装する。
- 研究開発項目 2：XR 会話 AI エージェントアプリケーションの開発：開発した通信インフラを活用したビデオ会議型、VR 型、AR 型のアプリケーションを開発する。
- 研究開発項目 3：実証実験：大規模なユーザ実験を通して通信速度・解像度による体験の比較を行う。

(2) 研究開発期間

令和 4 年度から令和 6 年度（3 年間）

(3) 受託者

株式会社エキュメノポリス <代表研究者>

(4) 研究開発予算（契約額）

令和 4 年度から令和 6 年度までの総額 300 百万円（令和 6 年度 100 百万円）

※百万円未満切り上げ

(5) 研究開発項目と担当

研究開発項目 1：スケールアウト可能な会話 AI エージェント・コンテナの開発
（エキュメノポリス）

研究開発項目 2：XR 会話 AI エージェントアプリケーションの開発（エキュメノポリス）

研究開発項目 3：実証実験（エキュメノポリス）

(6) 特許出願、外部発表等

		累計（件）	当該年度（件）
特許出願	国内出願	0	0
	外国出願	0	0
外部発表等	研究論文	0	0
	その他研究発表	11	7
	標準化提案・採択	0	0
	プレスリリース・報道	17	9
	展示会	4	2
	受賞・表彰	6	4

（7）具体的な実施内容と最終成果

研究開発項目2 XR 会話 AI エージェントアプリケーションの開発

ビデオ会議型言語学習支援エージェント「LANGX Speaking」を開発し、AI によるスピーキング能力の自動評価を実現した。本システムは、音声認識、言語理解、発話評価を統合した AI 技術を搭載し、学習者の発話内容をリアルタイムで分析し、CEFR（ヨーロッパ言語共通参照枠）基準に基づくスピーキングスコアを算出する。LANGX Speaking は、学習者の対話体験を向上させるため、適応型フィードバックシステムを採用し、発話内容の文法的正確性、流暢さ、発音の明瞭度などを詳細に分析する。このフィードバックは、学習者のスピーキング能力向上を支援し、効果的な英語学習を実現するための重要な要素となる。また、AI による自動評価機能を備えているため、評価の客観性が確保され、一貫性のあるスコアリングが可能となる。さらに、教師や教育機関の負担軽減にも寄与し、学習者一人ひとりに応じたパーソナライズされた学習支援を実現する。特に、教員が学習者のスピーキング評価を手作業で行う必要がなくなり、大規模な学習環境でも一貫性のある評価が可能となる点が大きな利点となっている。

研究開発項目 2-b) リッチクライアント型 言語学習支援エージェントサービスの開発

研究開発項目 1 で開発されてきたスケールアウト可能な会話 AI エージェント・コンテナ基盤の上に、リッチクライアント型の言語学習エージェントを開発し、WebGL 技術を活用したリアルタイムレンダリング技術を導入した。これにより、クライアント側での処理負荷を軽減しながら、没入感のある学習環境を提供可能とした。

研究開発項目 3 実証実験

本技術の有効性を評価するために、早稲田大学、九州大学、中央大学、立教大学、青山学院大学などの教育機関と連携し、累計 10 万回以上（2022 年度約 2 万回、2023 年度約 4 万回、2024 年度約 6 万回）の対話実証実験を実施した。その結果、システムの安定性やスケーラビリティが確認され、特に CEFR 準拠の自動スピーキング評価において、高精度なパフォーマンスを達成したことが明らかになった。

研究開発項目 3-b) 通信速度・解像度による体験の比較

ユーザー実験：本研究期間を通して、当初は通信速度や解像度が重要と想定していたが、実際には AI の反応速度や言語・非言語的な発話の適切性の方がより重要であることが明らかとなった。よって本研究では、同様の通信状況だった場合に、人間と AI エージェントの体験の違いを分析することで、システムの品質向上の知見を得ることとした。実験では、学習者が本研究で開発された AI エージェントおよびプロの人間チューターとそれぞれロールプレイおよび議論タスクを行い、対話体験の違いを比較した。実験の結果、人間チューターは、学習者の発話に対する適応的な応答や、表情・ジェスチャーを用いたコミュニケーションを通じて、より高いエンゲージメントとラポールを形成することが確認された。一方、AI エージェントは、一貫した評価基準を提供する点では有利であるが、対話の自然さや共感的なやり取りの面では人間チューターに劣る可能性がある。今後の課題として、AI エージェントの対話戦略の最適化が挙げられる。特に、ユーザーのエンゲージメントやラポールをリアルタイムで推定し、それに基づいて動的に対話スタイルを調整する技術の開発が求められる。また、異なる学習者層や対話タスクへの適用を通じて、評価指標の汎用性を検証することも重要である。

大規模な対話システムの継続的開発・運用管理の枠組みの検討と実践：生成 AI 技術の発展と対話システムの産業応用が進む中で、大規模な対話システムの開発・運用を効率的に管理するための効率的なワークフローが求められている。対話システムは、リアルタイム処理やマルチモジュール統合、品質評価の難しさ、ユーザーとのインタラクションを通じた学習の必要性といった独自の課題を持つ。既存のソフトウェア開発のパラダイムでは、こうした問題に十分に対応できていない。そこで、本研究では、対話システムの開発・運用を DevOps の CI/CD（継続的インテグレーション／継続的デリバリー）の枠組みと MLOps のデータ管理・モデル運用の概念を融合した「DialOps」を提案した。これにより、対話システムの開発から運用、改善のプロセスをシームレスに統合し、品質の維持と迅速な改善を可能にする。

DialOps の特徴として、第一に、対話システムの開発・運用フローを体系化し、各フェーズ（計画、開発、デリバリー、運用）の効率化を図った。第二に、品質保証のためのモニタリングを強化し、対話の品質を定量的に評価できる仕組みを構築した。特に、リアルタイム性の重要性を考慮し、応答時間の最適化やターンテイクの調整など、ユーザー体験を向上させる技術の開発を進めた。さらに、システムの外部要因（ネットワーク遅延、ユーザーの嗜好性、緊張度など）を考慮した評価指標を設計し、対話システムの実運用環境での継続的なパフォーマンス改善を可能にした。また、本研究で開発したエージェントを用いた実践を通じて、DialOps の有効性を検証した開発段階では、データの収集・管理を最適化し、発話ログの分析や対話のパフォーマンス向上に活用した。運用段階では、ユーザーのフィードバックをリアルタイムに反映し、対話システムのチューニングを継続的に行った。これにより、対話の流暢性やユーザーのエンゲージメント向上に寄与する結果が得られた。

（８）研究開発成果の展開・普及等に向けた計画・展望

本研究開発では、Beyond 5G 環境に適した高臨場感の会話 AI 技術の開発とその実証を完了し、研究開発の最終目標に対して極めて良好な進捗を達成した。特に、本研究を通じて、スケールアウト可能な会話 AI アーキテクチャを確立し、ビデオ会議型およびリッチクライアント型の AI エージェントの開発と実用化を成功させた。教育分野での技術適用を目的とした大規模な実証実験を実施し、その結果をもとに技術の有効性を検証した。国内の大学および中学・高校において、10 万回以上の対話実験を行い、AI エージェントの教育効果を評価した。これらの実験では、CEFR に基づくスピーキング評価の自動化が実現し、学習者の言語運用能力向上が確認された。また、エンゲージメントとラポールを指標とした評価手法の妥当性も実証された。さらに、ITU-T（国際電気通信連合）への標準化提案を進める準備を整え、グローバルな通信技術の発展に貢献する基盤を確立した。本研究の成果は、SIGDIAL 2024（Best Paper Award）などの国際学会で発表され、会話 AI 技術の新たな標準化に向けた基盤技術として評価された。今後は、ITU-T（国際電気通信連合）への標準化提案を進めるとともに、産業応用への展開を推進し、教育・ビジネス・医療分野への技術導入を加速させる計画である。

本研究の主な技術的成果として、スケールアウト可能な会話 AI エージェント・コンテナ開発を通して、大規模なユーザー同時接続に対応可能な分散型アーキテクチャが確立され、5G/Beyond 5G 環境下での低遅延かつ高品質な対話を実現した。また、高精細キャラクターレンダリングエンジンの開発を通して、リアルなアバター表現と非言語的コミュニケーションの精度が向上した。さらに、リッチクライアント型・シンククライアント型の通信アーキテクチャの統合を実施し、低スペックデバイスからハイエンド環境まで幅広いプラットフォームでの適用が可能となった。これらは、5G/Beyond 5G 通信環境下の汎用技術として転用性が高い。

本研究で開発した技術を元にして、当社は診断評価・学習支援向け会話 AI エージェントを活用し、対話を通じたスピーキング評価と学習支援を高度化することで、言語教育の DX（デジタルトランスフォーメーション）を推進するとともに、実社会で求められるコミュニケーション能力の向上を支援することを目的としている。本研究の成果も搭載される診断評価・学習支援向け会話 AI エージェントプラットフォームは、教育機関における言語教育の改革のみならず、対話型診断・学習支援技術を基盤として、企業の研修や専門職の対話能力向上、カスタマーサポートの品質管理といった多様な分野に応用できる可能性を持っている。本プラットフォームは、診断型会話 AI エージェントを大規模・高品質・高信頼・低コストで開発・運用できるプラットフォームである。本事業では、このプラットフォームを構築し、対話型診断への多様なニーズに対応したアプリケーション開発環境を提供することで、幅広い顧客への展開を可能とし、英会話に限らず、対話を通じた診断技術の革新を目指す。従来の言語評価システムでは、発話の正確性や流暢性の測定が主流であったが、本プラットフォームは、音声、テキスト、視線、表情、対話構造といったマルチモーダルデータを統合し、対話の流れや適切性を総合的に評価・診断する。