

# 音声対話ロボットによる文献調査補助システムの開発

## Development of a Literature Search Assistance System Using a Spoken Dialogue Robot

中浦 芳也<sup>1</sup> 早道 広峻<sup>1</sup> 松本 和幸<sup>1</sup> 篠山 学<sup>2</sup> 吉田 稔<sup>1</sup>

Y. Nakaura<sup>1</sup>, H. Hayamichi<sup>1</sup>, K. Matsumoto<sup>1</sup>, M. Shinozaki<sup>2</sup>, and M. Yoshida<sup>1</sup>

<sup>1</sup> 徳島大学

<sup>1</sup> Tokushima University

<sup>2</sup> 香川高等専門学校

<sup>2</sup> Kagawa Technical College

**Abstract:** 文献調査には多くの時間がかかり、専門知識がないと理解が困難である。また、卒業生の研究成果が十分に引き継がれないことが研究の継続性を阻害している。これらの課題を解決するため、音声認識技術と自然言語処理を用いた文献調査支援システムを提案する。提案システムの妥当性を評価するため、研究室に配属されたばかりの学生 10 名を被験者として、本システムを使用してもらい、アンケートによるシステムの機能面や妥当性の評価を行うための実験を行った。アンケートの結果から、提案システムの有用性と今後の課題が明らかとなった。本研究は、文献調査の効率化と研究の継続性向上に寄与すると考えられる。

## 1 はじめに

近年、膨大な文献が生成される一方で、必要な情報を効率的に取得することが困難になっている。文献調査には多くの時間がかかり、専門知識を持たない利用者にとっては内容の理解も難しく、これが研究や業務の効率低下の一因となっている。

また、大学や研究機関では、卒業生が残した研究成果や知見が十分に継承されず、特に高度に専門的な分野では、後輩が過去の研究内容を正確に把握・活用できないケースが少なくない。この問題は、継続的な研究の発展を阻害する要因となっている。

さらに、大規模言語モデル (LLM) の普及により自然言語処理を活用した情報検索技術は進化しているが、クラウドベースのシステムでは、機密情報の取り扱いに制約がある。一方、音声ベースのインターフェースは直感的な操作が可能で、特に手がふさがっている状況や迅速な操作が求められる環境で有用性が高い。

これらの背景から、効率的かつ継続可能な研究支援の仕組みが求められている。既存の AI を活用した論文検索システム (Perplexity AI[1], Elicit[2], Consensus[3]) は、自由記述のテキスト入力により文献検索を実現しているが、対話形式ではなく、またクラウドベースで運用されるため、内部の機密文献の取り扱いには不向きである。

本研究は、論文管理データベースを公開せず、ローカル環境で LLM を運用することで、システム全体をローカルで動作させるという新規なアプローチを採用している。これにより、機密情報の保護が強化されるとともに、利用者は音声対話ロボットを通じて直感的かつ効率的に文献調査を実施できる環境が提供される。具体的には、社会的対話ヒューマノイドロボット Sota[4]を用い、ロボットとの対話から興味のある分野のキーワードを抽出し、事前に研究室内で構築した論文要約データベースをキーワード検索する。検索結果は専用のチャットコミュニケーションツール上で表示され、利用者は過去の研究成果を容易に把握できる。

## 2 関連研究

### 2.1 オープンアクセスを見据えた研究室論文データベースの構築

宮代ら[5]は、研究室単位でのオープンアクセスを視野に入れた論文データベースの構築および運用について報告している。近年、学術成果の共有においてオープンアクセスの重要性が増している一方で、従来の学術成果は出版社が管理する有償リポジトリに依存しており、これが情報共有の制限や研究の発展を妨げる要因となっている。宮代らは、この問題

意識に基づき、研究室レベルでセルフアーカイブを実現する「研究室論文データベース」を構築し、一

般公開を行った。

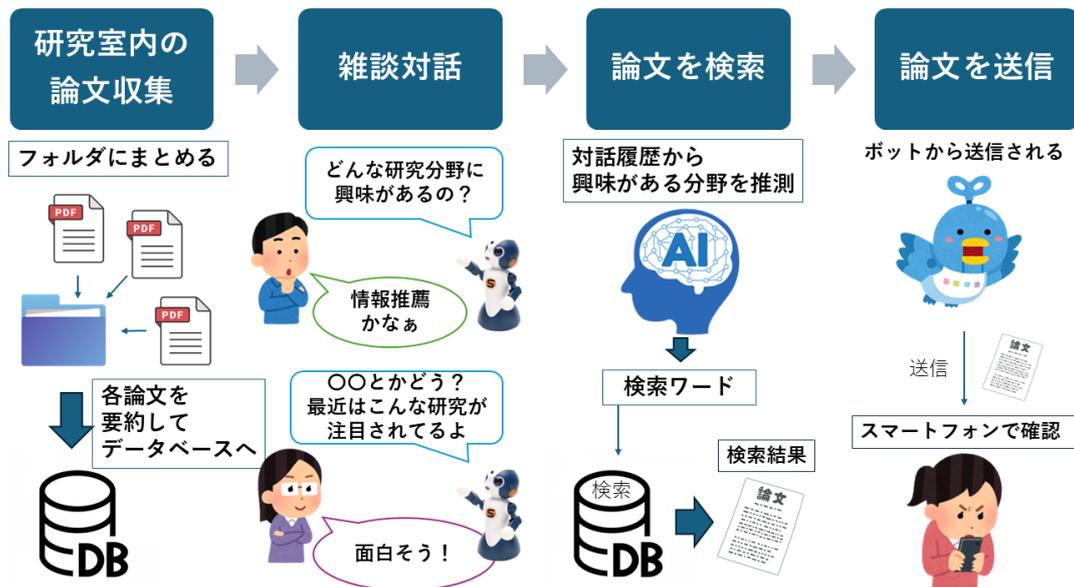


図 1: システムの流れ

本データベースは、既存の学術リポジトリのメタデータ仕様を踏襲するとともに、Google Scholar などの外部検索システムとの互換性を確保する設計となっており、研究室が保有する学術成果の広範な発信を可能としている。また、データベースの設計・運用に際しては、検索エンジン最適化 (SEO)、操作性の向上、並びにデータの継続的な更新・保守といった要素が考慮され、効率的かつ持続可能なシステムとして実現されている。

研究室レベルでのリポジトリ運用の利点としては、学術出版社に依存せず迅速に成果を公開できる点や、研究室固有の成果を広く認知させる手段として機能する点が挙げられる。一方で、運用には人的および技術的リソースの確保や、メタデータの標準化といった課題も指摘されている。

以上のような宮代らの取り組みは、オープンアクセスの促進に向けた実践的アプローチとして、学術成果の普及および活用の拡大に寄与すると考えられる。また、研究室単位での論文データベース構築という新たな視点を提供し、学術情報流通の未来像を示す試みである。

### 3 提案手法

提案手法は、ユーザとの対話を通じて必要な情報

を収集し、興味のある分野を推測して文献を提示するシステムである。まず、ユーザと対話を行い、音声認識や文章生成を通じて内容を処理する。その後、得られた情報をもとにユーザの興味を分析し、データベースから関連情報を検索する。検索結果はシステムからユーザに送信され、ユーザは提示された文献を確認することで、目的の資料を効率よく見つけることができる。ここから、図1のような流れで説明していく。



図 2: Sota の特徴

表 1: 論文の内訳

年度	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
卒論	8	10	8	7	9	10	6	10	9	12	9	6
修論	6	1	5	1	10	5	4	5	1	3	4	1

### 3.1 Sota について

本研究では対話ロボットとして Sota を用いる (図 2 参照). Sota にはカメラやマイクが搭載されており, 頭部や左右の腕である程度自由に動きをつけることができる. Sota には通常版とより高性能なマイクを搭載したインテリジェントマイク版の 2 種類があり, 本研究では後者を使用している. また, Sota 用に, Sota クラウドという音声認識と音声合成機能を提供するサービスがあるが, 本研究では使用しない.

### 3.2 論文データベースの構築

本研究では, 過去 10 年間にわたる研究室の学士論文および修士論文を対象に, それぞれの論文の要約データベースを構築した (表 1 に各論文数を示す). 具体的には, 各フォルダに保存された PDF ファイルを一括処理し, 前処理を通じてテキストデータを抽出した. 抽出したテキストに対して大規模言語モデル (LLM) を適用し, 自動要約を実施した後, その結果をデータベースに格納することで, 効率的な情報検索および閲覧が可能な環境を実現した.

本研究では, 要約データを作成する際に, 一定のフォーマットに基づいたプロンプトを使用した. 具体的には, 以下のプロンプトを用い, 論文本文を入力として要約を生成した.

プロンプト:

以下の論文を日本語で要約し, 指定されたフォーマットに従って出力してください.

フォーマット:

1. タイトル: [論文のタイトル]
2. 著者と発表年: [著者名, 発表年]
3. 目的: [研究の目的を 1~2 文で]
4. 方法: [研究手法やアプローチを 1~2 文で]
5. 結果: [主要な結果を簡潔に]
6. 結論と意義: [結論とその意義を簡潔に]
7. キーワード: [関連するキーワードを 3~5 個]

論文内容:

[要約対象の論文本文]

本プロンプトでは, [要約対象の論文本文] の部分に実際の論文本文を入力し, 指定されたフォーマットに従って要約を生成した. このプロンプトの採用により, 要約の統一性を確保し, 異なる論文間での比較が容易となった.

さらに, 本研究では, 生成された要約の圧縮率についても調査を行った. その結果, 元の論文の文字数に対する要約文の文字数の平均比率は 1.39% であり, 大幅に圧縮されていることが確認された. この圧縮率の結果は, 要約の簡潔さを示す一方で, 情報の欠落の可能性についても考慮する必要がある.

### 3.3 雑談対話

本システムでは, まず Sota が利用者に対して「どのような研究に興味があるか」を尋ねることで対話を開始する. 利用者は自身の興味に関する回答を行い, その内容に基づいて Sota は雑談形式の対話を展開する. この対話は単なる情報収集に留まらず, 利用者と Sota の双方の発言がすべて記録され, 後続の論文検索のための重要なデータとしてデータベースに蓄積される. 複数回にわたる対話を通じて, 利用者の関心や研究ニーズが明確化された後, システムはその対話内容を活用して論文検索フェーズへと移行する. このプロセスにより, 利用者は自然な会話を通じて, 自身に最適な文献情報にアクセスできることが期待される.

また, セキュリティおよびプライバシー保護の観点から, 音声認識, 文章生成, 音声合成の各処理をすべてローカル環境で実装している. 音声認識では, まず WebRTC VAD[6]を用いて入力音声から発話を検出し, 発話が認識されると自動的に録音を開始する. その後, 録音された音声は ReazonSpeech[7]によりテキストデータへ変換される. 対話や検索クエリに応じた自然な文章生成には, 日本語に特化して学習された Llama3[8]モデル Llama-3-ELYZA-JP-8B[9]

を採用しており、生成されたテキストは Style-Bert-VITS2[10]によって音声に変換される。この音声合成モデルは、独自に学習したモデルを用いることで、自然で聴きやすい音声出力を実現している。以上の各技術をすべてローカルで運用することにより、外部クラウドへのデータ送信を回避し、機密情報の漏洩リスクを大幅に低減するとともに、システム全体のセキュリティおよび応答速度の向上を実現している。図3に対話システムの概要を示す。

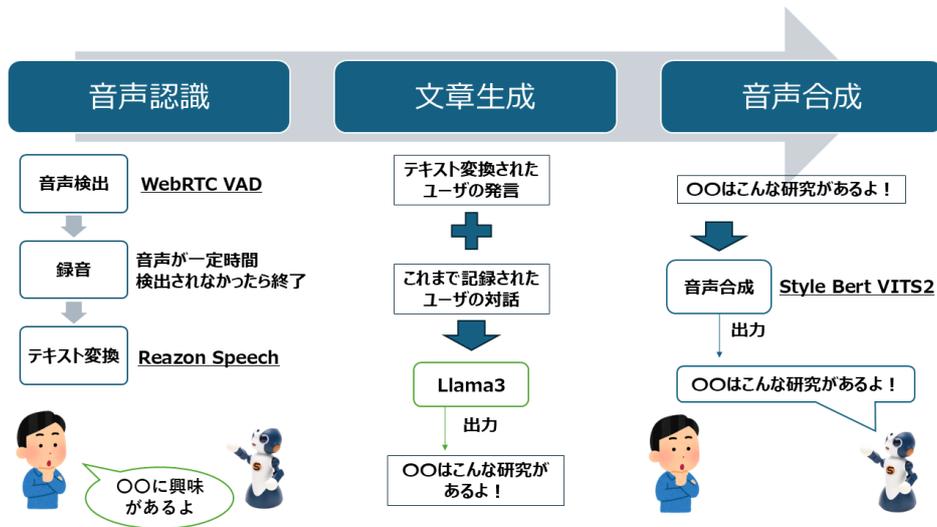


図 3: 対話システムの概要

トル、著者、発表年、及び要約が含まれている。検索処理では、抽出されたキーワードを基にタイトルに対する部分一致の評価を行い、関連度の高い上位3件の論文を抽出して利用者に提示する。これにより、利用者は対話を通じて得られたキーワードに沿った、関連性の高い文献情報に迅速にアクセスできる。なお、本検索プロセスもシステム全体と同様にローカル環境で実装されているため、外部へのデータ送信が行われず、機密情報の保護と高い応答速度が確保される。

### 3.5 論文の送信

本システムでは、検索によって得られた上位3件の論文に関する情報を利用者に提示する。具体的には、該当論文のタイトルおよび要約をコミュニケーションツールを通じて送信し、利用者が詳細を確認できるようにする。この際、検索結果の一覧を視覚的に整理し、論文ごとの概要を即座に把握できるインターフェースを設計することで、利便性を向上させる。また、検索結果の内容は、対話を通じて抽出

### 3.4 論文の検索

論文検索機能は、本システムの中核的な要素の一つであり、利用者との対話から抽出されたキーワードに基づいて、関連文献を迅速に抽出することを目的としている。システムは、Sota との対話の中で利用者の興味・関心を表すキーワードを抽出し、そのキーワードを用いて、事前に構築された論文要約データベースに対してタイトル検索を実施する。データベースには、過去10年間にわたる研究室の論文が要約された情報が格納されており、各論文にはタイ

されたキーワードとの関連性に基づいて並べ替えを行う。さらに、利用者が論文の全文を確認できるように、検索結果表示の後に対応する論文のPDFファイルを送信する。これにより、利用者は興味を持った論文を迅速に閲覧し、研究に活用できる。

## 4 実験

本実験では、システムの有用性および利便性を評価するため、10名の被験者に協力を依頼し、実際にロボットとの対話を行ってもらった。対話終了後、被験者にはアンケートに回答してもらい、システムの評価を実施した。アンケートでは、本システムの有用性や継続利用の意向、論文要約の理解しやすさ、関心度などについて、5段階評価で回答を求めた。また、自由記述形式の質問を設け、被験者の関心分野や本システムに対する改善点についての意見を収集した。

### 4.1 実験の準備

本実験では、被験者ごとの対話履歴を適切に管理

するため、対話開始前に QR コードを用いたユーザ認識を行う。QR コードの発行には、実験で使用するコミュニケーションツール上に事前に作成したサーバを利用する。被験者は、実験開始前にこのサーバに参加し、サーバ内のチャット Bot を通じて自身の QR コードを取得する。この QR コードをシステムに提示することで、被験者ごとの対話履歴を識別し、適切に管理することが可能となる。

## 4.2 結果

システムの有用性や推測精度を評価するために、アンケート調査および対話ログの分析を行った。本章では、これらの結果について報告する。まず、アンケート結果をもとにシステムの利用意向や改善点について述べる。次に、音声対話を通じた興味分野の推測精度について分析し、推測が成功したケースと失敗したケースの特徴を明らかにする。

### 4.2.1 アンケート結果

アンケート結果を表 2 に示す。本実験に参加した 10 名の被験者からのアンケート結果に基づき、システムの有用性および利用意向について評価した。Q1～Q5 の項目に関しては、いずれも平均 4 点前後の高評価を得ており、システムが有用であると感じられていることが確認された。特に、Q7（コミュニケーションツールとの連携について）は 4.7 点と高評価を得ており、操作性や利便性の向上に寄与していると考えられる。

一方で、Q6（引き継ぎ時にこのシステムがどれくらい役立つか）については、平均 3.4 点と比較的低い評価となり、システムが引き継ぎ作業においてどれほどの支援ができるかについて、さらなる改善の余地が示唆された。

自由記述では、システムの音声インターフェースや迅速な検索結果に対する好意的な意見が多く寄せられたが、引き継ぎ機能や検索結果の精度に関しては改善が求められた。特に、検索精度や推薦理由の明示、引き継ぎ時のサポート機能に対するニーズが高く、これらはシステムの有用性を向上させるための重要な改善点として挙げられた。

### 4.2.2 推測精度の分析

対話ログ分析の結果、音声対話を通じたユーザの興味分野推測において、推測が成功したケースと失敗したケースには明確な違いがあることが分かった。推測が成功した対話では、ユーザの発話が比較的

長く、研究テーマや興味分野について積極的に言及していた。一方、推測が失敗した対話では、ユーザの発話が短く、曖昧な返答が多かった。また、アシスタント（LLM）の発話内容も推測精度に影響を与え、推測が成功した場合は具体的で詳細な発話が行われ、ユーザの関心を引き出す効果があった。しかし、推測が失敗した場合は、アシスタントの発話が一般的で曖昧であり、ユーザの関心を正確に捉えられなかった。

表 2: アンケート結果

番号	質問内容	平均
Q1	使ってみて楽しかったですか？	4.1
Q2	役に立ちそうでしたか？	4.1
Q3	今後も使ってみたいですか？	3.9
Q4	送られてきた論文の要約はわかりやすかったですか？	4.0
Q5	論文の内容は興味ありましたか？	4.0
Q6	もし引き継ぎを行うとなった場合このシステムでどのくらい担えると思いましたか？	3.4
Q7	コミュニケーションツールとの連携についてどのくらい便利だと感じましたか？	4.7

## 5 考察

本実験の結果、システムは全体的に好評を得たが、いくつかの改善点も明確になった。まず、システムの発話生成スピードについて、一部の被験者から改善の余地があるとの指摘があった。生成速度を向上させることが、システムの使用感の向上に寄与する可能性が高い。また、研究テーマの調査効率を向上させるために、Sota 側から研究提案を行う機能を追加すべきだという意見があった。こうした機能の追加により、ユーザはより効率的に研究を進めることができると考えられる。

次に、Sota の動きに関しては、親しみやすさを感じる一方で、動きが過剰になると注意が散漫になり、内容の理解が難しくなるとの意見があった。このため、動作の頻度やタイミングに関して調整が必要であると考えられる。また、会話の終了時に不自然さを感じるユーザもおり、適切な終了の挨拶を加えることにより、対話の流れをより自然にすることができると考えられる。

さらに、システムによる研究テーマの推薦について、推薦理由を示すことが有益であるとの意見があった。推薦理由が明確であれば、ユーザは推薦され

た研究を納得して受け入れることができ、システムの信頼性向上に繋がると考えられる。加えて、自由記述においては、システムのモーションが途中で停止したり、英語表記の読み上げに関する不満が挙げられた。特に、モーションが終了間際に止まることや、英語表記の音声合成処理に不便さを感じたユーザがいたため、これらの問題に対する改善が求められる。

これらの結果を踏まえ、システムは全体として有用であるとの評価を得たが、今後の改善点として、生成スピードの向上、動作の調整、推薦理由の提示、安定した動作の実現が挙げられる。

システムの推測精度に関しては、ユーザの発話内容や長さが大きな影響を与えていることが分かった。短い発話でも適切に推測できるアルゴリズムの改善が必要である。また、アシスタントの発話内容が具体的であるほど、推測精度が高くなることが示された。今後は、ユーザの発話から意図をより正確に汲み取るための対話アルゴリズムの改良が重要である。

本研究とは別に、元の論文の文字数に対する要約文の文字数の比率を調査した結果、平均 1.39%であった。これは、元の論文と比較して要約文が極めて短縮されていることを示している。

## 6 まとめ

本研究では、音声対話ロボットを用いた文献調査補助システムを開発した。本システムは、ローカル環境で動作する音声対話を通じて、ユーザが関心を持っている研究分野を推測し、推測された分野に基づいてデータベースから関連する資料を検索し、さらにコミュニケーションサービスでその結果を共有する機能を持つ。これにより、研究に不慣れなユーザに対して文献調査の支援を行うことを目指した。

10名の被験者による対話実験を実施し、アンケートによってシステムの有用性と改善点を評価した。実験結果から、本システムはユーザの関心分野をある程度推測できることが確認されたが、完全には一致しないケースもあり、推測精度の向上が課題として浮き彫りになった。さらに、ユーザの意見を分析したところ、対話の自然さ、システムの動作、情報提示の方法という3つの観点で改善の余地があることが分かった。一方で、論文検索や要約の迅速性については肯定的な意見が多く、システムの基本機能における有用性が確認された。

今後の課題として、まずユーザの興味推定精度の向上が挙げられる。現状ではユーザの興味と推測された研究分野には関連性が見られるものの、一致しないケースも存在するため、個々のユーザの発話履歴や過去の興味傾向を考慮した精度向上が求められる。次に、対話の自然さの向上が必要である。実験

結果から、対話の終わり方や質問のタイミングに関する改善点が指摘されたため、対話のアルゴリズムを改善し、会話の流れをスムーズにすることが重要である。また、ユーザの発話意図をより適切に理解し、自然な応答を生成するための対話制御の最適化も必要である。さらに、情報提示手法の改善も重要な課題である。システム側から能動的に情報を提示する機能や、推薦理由を明示する仕組みが求められている。

加えて、データのセキュリティ確保の観点も重要である。本システムでは検索結果を共有する手段として Discord[11]を利用しているが、クラウドベースのサービスであるため、機密情報の漏洩リスクが懸念される。今後は、オンプレミス環境での代替手段の導入や、通信内容の暗号化、アクセス制御の強化などを検討する必要がある。

今後は、ユーザが納得しやすい形で情報を提示し、対話のリアルタイム性を保ちながら、迅速に文献情報を提供できる文献調査補助システムの構築を目指していく。

## 謝辞

本研究の一部は JSPS 科研費 JP21K12141, JP22K12213, JP24K15193 の助成を受けたものです。

## 参考文献

- [1] Perplexity AI: <https://www.perplexity.ai>
- [2] Elicit: <https://elicit.com>
- [3] Consensus: <https://consensus.app>
- [4] コミュニケーションロボット「Sota (ソータ)」 | ヴイストン株式会社
- [5] 宮代 理弘, 宮下 芳明: オープンアクセスを見据えた研究室論文データベースの構築, 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2017-HCI-174, No. 12, pp. 1-7, (2017)
- [6] WebRTC VAD: <https://github.com/wiseman/py-webrtcvad>
- [7] ReazonSpeech: <https://research.reazon.jp/projects/ReazonSpeech/index.html>
- [8] Llama: <https://www.llama.com>
- [9] Llama-3-ELYZA-JP-8B: <https://huggingface.co/elyza/Llama-3-ELYZA-JP-8B>
- [10] Style-Bert-VITS2: <https://github.com/litagin02/Style-Bert-VITS2>
- [11] Discord: <https://discord.com>