

# インフラ領域における職人の技の 伝承教育システム

国立研究開発法人理化学研究所

- 職人の技を科学的に分析した過程を「工学的前兆判定 AI」として確立し、その基盤技術を組み込んだ、「職人教育システム」によって、技の習熟度の可視化に成功し、生産性向上の見込みが立つことを確認した。
- 「工学的前兆判定 AI」を組み込みソフトウェアとして実装したロボット検査システムを構築し、業務訓練や支援に役立てるため、業務状況の計測や評価をもとにしてタスクデータ・情報ベースを構築した。

## 1 研究の目的

日本のインフラを支えている優秀な職人の人材不足や、高齢化に対応することを目的に、職人の技を科学的に分析。その判断過程を、AIで再構築した基盤技術「工学的前兆判定 AI」として確立した。この基盤技術をソフトウェアとして組み込んだ、「職人教育システム」を構築するとともに、SIP 第1期で開発・改良されたインフラ計測デバイス（MIMM およびレーザー打音）に対して、「工学的前兆判定 AI」をデバイス組み込みソフトウェアとして実装したロボット検査システムを構築した。「職人教育システム」を用いて訓練を受けた職人の最終判断と、ロボット検査システムによる支援システムとの連携により、Society 5.0において実現を目指すインフラ検査のCPS\*1（Cyber Physical System）型社会実装形態を実現した。

## 2 実施期間と方法

### (1) 実施期間

2018年度～2022年度

### (2) 実施方法

認知的インタラクションによる職人技の研究を基本に、インフラ領域における職人技の伝承教育と機械化（ロボット化）について、全体構造の出口（社会実装）に向かって各研究テーマを遂行し、研究のフローを作成した（図1）。本研究は、図1において「人による診断」で示されたフローに対応し、以下の2点の開発を行った。

#### 1) インフラメンテナンスにおける職人の技の解析および体系化に関する研究開発（理化学研究所）

作業動画収集のためのウェアラブルカメラシステム開発と、作業動画から技データを生成、処理するソフト開発によって本研究を進めた（両者を合わせて職人技データ収集システムと呼ぶ）。前者では、作業員への負荷を極力増やさずに動画収集するシステム開発を進めた。後者では、作業観察と聞き取りによる技の要素の特定のうち、それらのデータ化（技データ）と、技データから人の判断要素の抽出を実現するアルゴリズム開発を進めた。

#### 2) 科学的手法に裏付けられたインフラメンテナンス教育システムの開発（理化学研究所）

マシン（教育システム）による技術習得の手法確立と、

\*1 現実世界でセンサーが収集した情報をサイバー空間で活用して解析し、経験や勘ではなく定量的な分析でさまざまな産業に役立てる取り組み。

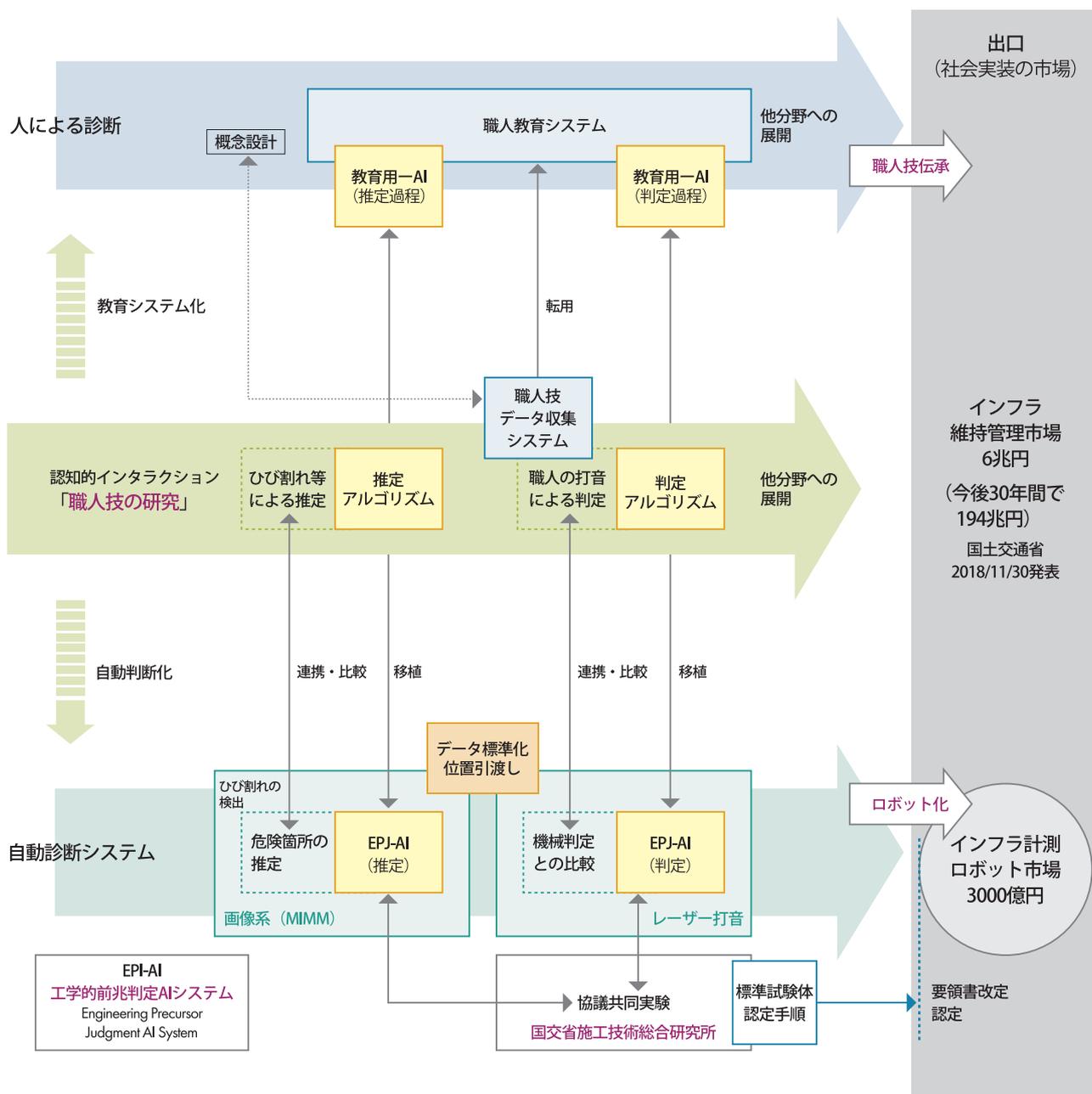


図1 研究のフローチャート

それを実現するシステムの開発によって本研究を進めた。前者については、実際の教育方法の聞き取りや現場での教育体験を通して進めた。後者については、1) による職人技データ収集システムに、技データと判断要素を可視化する技術の導入など、教育効果（生産性）を上げる機能を追加導入する方針で進めた。また、システムの使用方法や効果、応用については、点検員からの聞き取り調査をもとに改善していく方法をとった。

### 3 本研究の成果

#### (1) インフラメンテナンスにおける職人の技の解析および体系化に関する研究開発（理化学研究所）

以下の三つの成果を挙げた。

- 訓練環境（室内で供試体を使用）にて、打音点検作業から職人技の要素（探索の技、絞込み・推定の技、音聞き分けの技）を収集するウェアラブルカメラと、モバイルPCからなるシステムの実証機を完成させた。

- 実際のトンネル現場における打音点検作業から職人技を収集するウェアラブルカメラと、モバイル PC からなるシステムについて、実証機を完成させた（図 2）。
- 職人技のデータ化と、そこから点検員による判断過程と判断基準を抽出するソフトウェア（Eyelex）を完成

させた。

## (2) 科学的手法に裏付けられたインフラメンテナンス教育システムの開発（理化学研究所）

ウェアラブルカメラと、ソフトウェア（EyelexTablet）



図 2 収集した作業データ  
プロットがハンマー打撃点、紫領域が点検員が判断した異常領域に対応、各プロットごとに打撃音が付属



作業



処理結果確認

図 3 教育システム使用の様子

ウェアラブルカメラ（左図矢印）で撮った作業動画がその場でソフトウェア EyelexTablet を搭載したモバイル PC（右図矢印）に転送、処理される

を搭載したモバイル PC からなる、教育システムを完成させた(図3)。点検の技である、探索力、絞込み・推定力、音判断力の3要素を可視化するとともに、それらによる総合判断力をスコアとして表示した。その結果、新人と熟練者の技(習得度)の違いを、可視化イメージとスコアによって比較することが可能となった(図4)。教育訓練では、ウェアラブルカメラ付きヘルメットを装着して作業を実施後、その場でスコアを確認できるとともに、減点の理由を三つの要素技のスコアから理解し、技術習得を一人で進めることができる手法を開発した。

なお、本システムの生産性向上は以下で見積もることができる。変状判定にて、熟練者が教授する項目を、打音とひび割れ構造の相関性、環境、経年、施工法と仮定し、教育システムでアシストする部分を「打音とひび割れ構造の相関性」の一部とする。それが、教授する項目の中で占める時間的な割合は40%と見込めるため、教育システムは職人1人の労力に対して0.4人分の労力に相当する。本システムを従来の教育と併用することで等価的に熟練者は1.4人分の労力を有することになり、職人の

2年の教育期間に対して熟練者を1.4人分当てることができるため、教育期間は1.4年(2年/1.4人分)に短縮することができる。これから教育システムを導入することで生産性は約143%(2年/1.4年)と見積もられる。

## 4 まとめと今後の展望

被打面画像上への打撃点と付随するハンマー打音のデジタルデータ化の技術や、それに基づく、熟練者の打音点検の技と判断過程および基準を可視化する技術を開発した。これにより、熟練者からの技の習得について、これまで困難であった生徒の自習・復習学習が可能になったほか、学習の場が得られる生徒の数に制限がなくなった。その結果、効率的な技術習得の訓練を実現でき、点検員の育成促進に期待できる。

この技術は、点検業務の支援に応用できるだけでなく、打音に頼るほかの点検用途にも応用が期待できる。

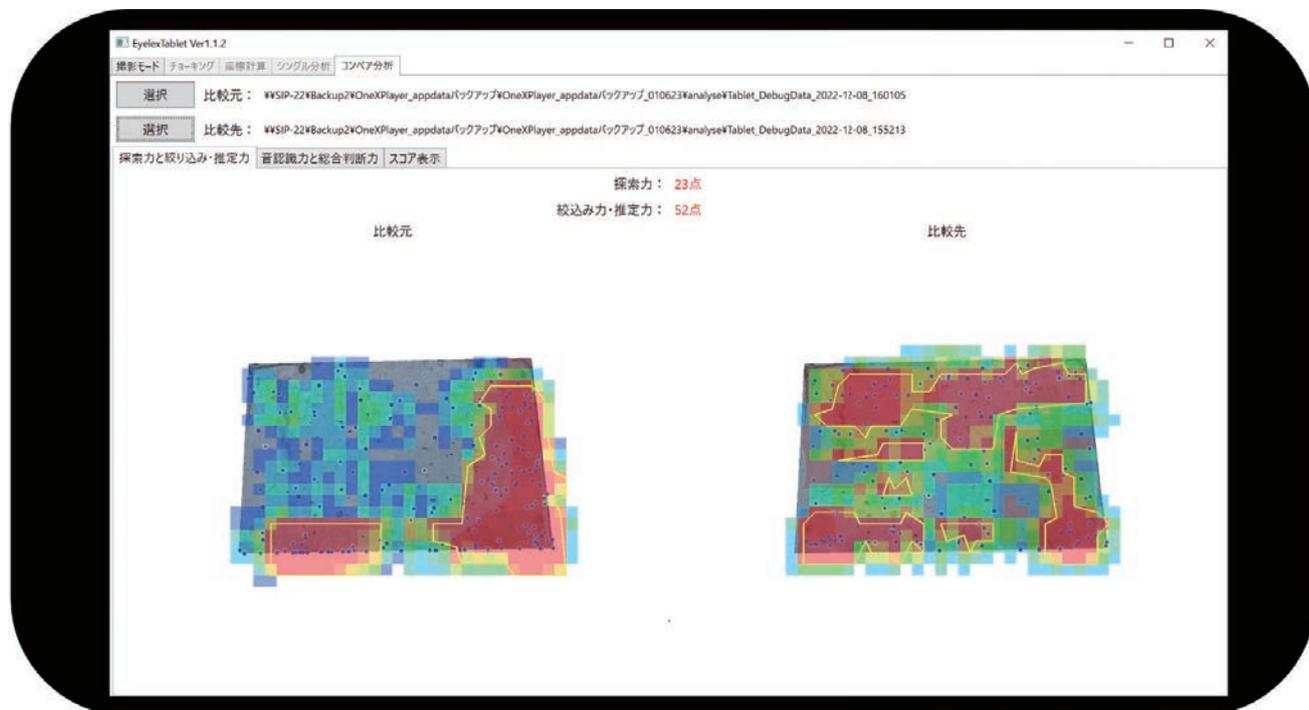


図4 タブレット PC に表示された使用者(左)と熟練職人(右)の技(探索力と絞込み・推定力)の比較の例  
 探索力は黄色線領域で、絞込み・推定力は色と濃淡で、スコアは赤字で表示されている

# インフラ領域における打音検査の機械化の推進と社会実装

計測検査株式会社、国立研究開発法人量子科学技術研究開発機構、株式会社フォトンラボ、一般社団法人社会基盤情報流通推進協議会

- 「工学的前兆判定 AI」を組込みソフトウェアとして実装したロボット検査システムを構築し、業務訓練や支援に役立てるため、業務状況の計測や評価をもとにタスクデータ・情報ベースを構築した。供用中の道路トンネルを対象とした本 AI の検証試験では、検査員による健全度判定を 90%以上の精度で再現することに成功した。加えて、データ解析時間の短縮により、従来と比較して約 190%への生産性の向上が見込まれる。本 AI で判定した結果をデータベースに登録することで、複数の検査装置の連携を容易とし、かつデジタルデータとして他者との情報共有を可能とした。
- 研究参画機関から研究成果を引き継ぎ、一連の機器システムの計測サービスは計測検査の画像計測サービスと連携して顧客に引き渡されるようにした。

## 1 研究の目的

日本のインフラを支えている優秀な職人の人材不足や、高齢化に対応することを目的に、職人の技を科学的に分析。その判断過程を、AI で再構築した基盤技術「工学的前兆判定 AI」として確立する。この基盤技術をソフトウェアとして組み込んだ、「職人教育システム」を構築するとともに、SIP 第 1 期で開発・改良されたインフラ計測デバイス（MIMM およびレーザー打音）に対して、「工学的前兆判定 AI」をデバイス組込みソフトウェアとして実装したロボット検査システムを構築。「職人教育システム」を用いて訓練を受けた職人の最終判断と、ロボット検査システムによる支援システムとの連携により、Society 5.0 において実現を目指すインフラ検査の CPS\*1（Cyber Physical System）型社会実装形態を実現する。

## 2 実施期間と方法

### (1) 実施期間

2018 年度～2022 年度

### (2) 実施方法

認知的インタラクションによる職人技の研究を基本に、インフラ領域における職人技の伝承教育と機械化（ロボット化）について、全体構造の出口（社会実装）に向かって各研究テーマを遂行して研究のフローを作成した（図 1）。本研究は図 1 において「職人技の研究」「自動診断システム」に示される。

本研究では、トンネルの点検で行われる打音検査の機械化を推進するために、図 2、図 3 に示す 2 つの計測デバイス（MIMM、レーザー打音）から得られるデータを診断するための AI を開発するとともに、AI により判定したデータを登録し共有するためのデータベースを開発した。通常の打音検査では、検査員がトンネル内壁（覆

\*1 現実世界でセンサーが収集した情報をサイバー空間で活用して解析し、経験や勘ではなく定量的な分析でさまざまな産業に役立てる取り組み。

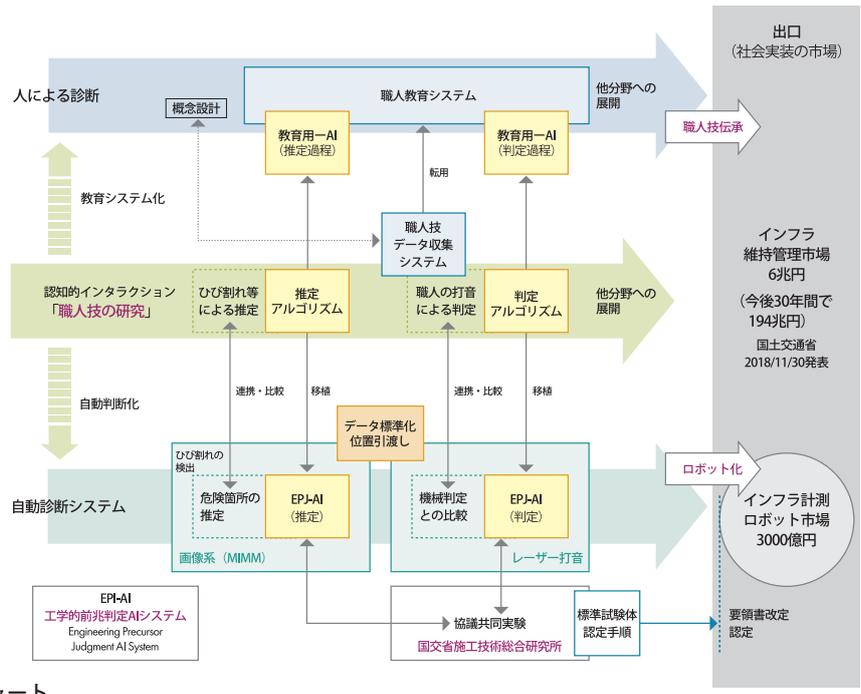


図1 研究のフローチャート



図2 Mobile Imaging technology & Mobile Mapping system (MIMM)

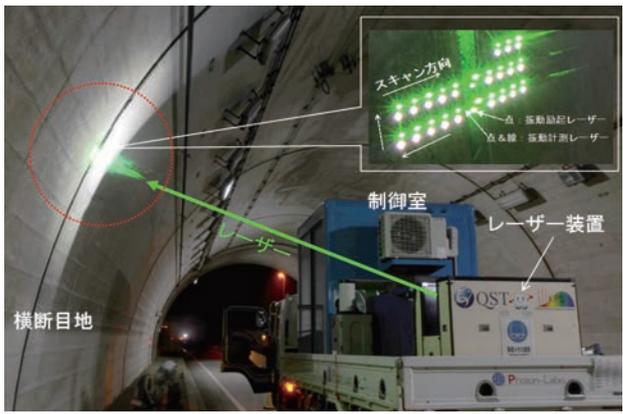


図3 レーザー打音検査装置

工面) を目視で確認し、点検が必要と判断した箇所(ひび等がある)をハンマーで叩いて内部に空洞などの欠陥が無いかを診断している。MIMMは、走行しながらトンネル覆工面の写真を撮影するデバイスであり、本研究で開発したAI(図1のEPI-AI)により点検が必要な箇所を抽出し、レーザー打音に指示を出す「検査員の目」としての働きをする。レーザー打音は、ハンマーの代わりに近赤外線のパルスレーザーをコンクリート表面に照射することで振動を励起し、その振動を「検査員の耳」に相当するレーザードップラー振動計測により計測し、EPI-AIによりコンクリートの健全度を判定する。

1) 画像計測による危険箇所の高精度抽出に関する研究

**開発 (計測検査株式会社)**  
 近年、一部の道路定期点検では、走行型ロボット撮影技術を用いて人による画像判断を行っている。その画像の危険箇所推定精度・速度の向上を図った。職人技の集積である点検要領記載の基準のうち、ひび割れ閉口と閉口しそうな箇所、また、ひび割れ以外(漏水、遊離石灰、ジャンカ、補修跡、前回浮き箇所のチョーキング囲い)を判定する画像ベース判定技術を確認した。  
 正解データに類似する、誤検出の影響がある学習データを見直し、約10万個のサンプルによる学習を行った。画像計測システムで撮影取得されたトンネルの撮影展開画像に発生している漏水や、遊離石灰などの劣化箇所をMASK-RCNN\*2ベースAIを用いて自動抽出する技術を

\*2 入力された画像の中から、あらかじめ指定された種類の物体を認識してその位置などを特定するアルゴリズム。

構築した。

速度については、現状手法の劣化を画像から人カトレースする作業をAI抽出で支援することで作業速度の向上を図った。

最終年度では、実際の道路トンネルを走行型トンネル点検専用車輛（MIMM）で計測し、レーザー打音システムとの連携用に、AI抽出された危険箇所位置データなどの各種メタデータを作成してデータベースへ登録。ひび割れとそのほか変状の抽出から得られる、複合変状の危険性推定を行った。

## 2) レーザー打音によるコンクリート内部欠陥の高精度判断に関する研究開発（量子科学技術研究開発機構）

レーザー打音検査では、パルスレーザーの照射により発生したコンクリート表面の微弱な振動を別のレーザーを用いて遠隔から計測し、その振動を解析することで、欠陥の検知を行う。コンクリート内部にうき等の欠陥がある場合、内部に伝搬した振動が欠陥位置で反射・干渉することで共鳴振動が生じ、特定の周波数成分が強く現れる（卓越）ため、この周波数の卓越から欠陥を検知する。この欠陥検知の原理は通常の打音検査と同様であり、高い互換性が期待できるが、本研究以前では、欠陥の有無のみの判定であり、また機械化特有の信号ノイズに起因するスペクトルの卓越による誤判定といった課題があった。本研究では、国土交通省の道路トンネル点検要領に準拠した多段階（Ⅰ：健全、Ⅱ：要観察、Ⅲ：修繕推奨）の健全度にノイズ判定を加えた4段階の判定に対応するAIを開発した。検査員の判定を再現するAIの開発で重要となるのは、実構造物に発生する欠陥の構造を再現しつつ欠陥の構造が明らかであり、かつ検査員により健全度が評価された供試体を用いて教師データを蓄積することである。コンクリート工学の専門家および現役の検査員と共同で新たに供試体を開発し、これを用いてAIの学習用に振動データを蓄積し、現役の検査員による判定を90%以上の精度で再現することを目標とした。開発したAIを実際に使用されている道路トンネルから得られた振動データに適用し、検査員の判定と比較することで判定精度の検証を行った。

## 3) インフラメンテナンスのためのデータベースの相互連携プラットフォームに関する研究開発（社会基盤情報流通推進協議会）

2020年度のモックアップや仕様案をもとに、画像計

測データとレーザー打音データを連携するデータプラットフォームのプロトタイプを構築し、サンプルデータをもとに動作確認した。また、ユーザーインターフェースや画面表示する内容などを改善するとともに、将来的な実運用を見据えての課題および対応案について整理した。対応案として、レーザー打音時に危険箇所の計測優先度を、トンネル単位に加えてスパン単位でも提示できるようにする、トンネル内の通信環境が悪い場所でも利用できるようオフラインでの利用を可能にする、などを示した。また、プロトタイプ構築にあわせ、データベース仕様書案の改定およびデータ流通モデルアーキテクチャ案を作成した。

## 4) 工学的前兆判定AIを搭載したインフラ検査システム等の製品化と現場適用（株式会社フォトンラボ：実施期間 2020年度～2022年度）

基礎技術の研究と社会実装の両立のため、2017年8月（SIP第1期の第4年度）に理研ベンチャー「フォトン応用計測研究所」を設立。後に、株式会社フォトンラボに社名変更した。

株式会社フォトンラボでは、道路トンネルの現場で実証実験や社会実装を行い、性能限界や課題を抽出して改良の方向性を見出し開発チームにフィードバックした。レーザー打音検査装置を8か所の現場に持ち込み、うき検出の限界事例「目地部および既補修箇所」を研究者とともに取得して課題解決について協議した。画像計測に関しては、現場からの画像データを対象に、ひび割れ抽出する研究開発の課題解決について、レーザー打音との連携運用による社会実装サイドから協議した。

教育システムについては、プロトタイプを1か所の現場に持ち込み、現場専門家が使用しながら課題としての「打点の視界はずれ」を発見し、開発チームにフィードバックした。新しい用途として、「専門家同士の協議用データ収集・解析装置」を提案した。

出口戦略として、研究参画機関から研究成果を引き継ぎ、各市場に向けて、関係する機関・企業と綿密な戦略構想の下で提携・協力関係を構築。一連の機器システムの計測サービスは、計測検査の画像計測サービスと連携して顧客に引き渡されるようにした（図4）。

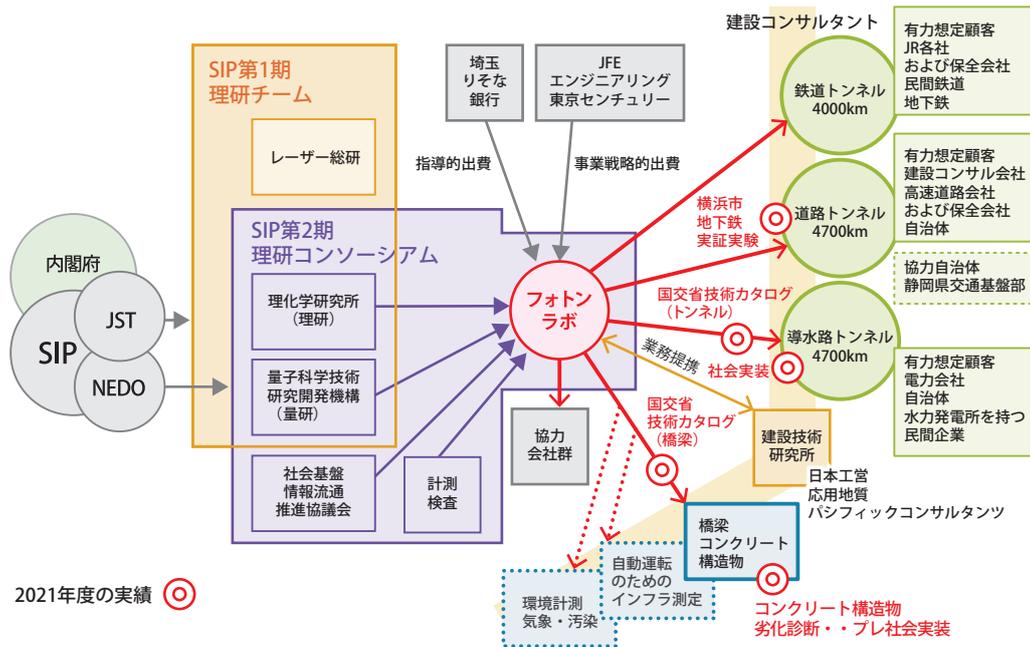


図4 全体戦略図

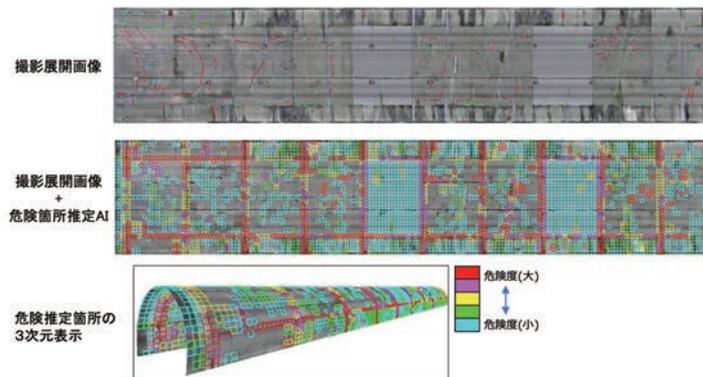


図5 覆工展開画像から、危険箇所推定 AI により危険箇所を抽出

### 3 本研究の成果

#### (1) 画像計測による危険箇所の高精度抽出に関する研究開発 (計測検査株式会社)

##### 1) ひび割れと、そのほかの変状のAI 自動抽出

走行型トンネル点検専用車輛 (MIMM) で得られる撮影展開画像から、ひび割れは、U-Ne\*3 ベースの AI 抽出手法を用いて独自の線検出アルゴリズムを構築した。図5に、展開画像、AIによる危険箇所の抽出、危険箇所の3次元表示の例を示す。画像から劣化を人力でトレースする作業を100%として、AI抽出が関与するのがその80%に相当するためAI再現抽出率の目標値を5%の振れ幅をもたせて75~80%に設定し、結果として

76.7%の再現抽出率が得られた。その他変状はMASK-RCNN ベースの AI 抽出手法を用いて領域分割アルゴリズムを構築し、71%のAI再現率を得られた。目標に対して未達であったが、欠損箇所の一部画像が不鮮明な箇所があったため、不鮮明な箇所をいかに取り除けるかが課題であることが分かった。積み上げにて人での変状箇所の検出する時間12.3時間に対して、AIによる自動抽出を加えることで検出時間が6.75時間に短縮され、生産性として約180%向上することが分かった。

##### 2) 危険箇所の判定技術

ひび割れとその他の変状のAI抽出結果をもとに、危険箇所を推定した。矩形メッシュ内において複数発生する、劣化箇所の位置や属性による重み付けを行い、加

\*3 生物医学のために開発された、深層学習のアルゴリズム。

点式判定による危険性の優先順位付けを行った。浮き・剥離などの危険箇所の人判定結果に対して、AIの判定結果一致率は78.3%であった。不一致の要因の一つとして人によっても危険の判断に迷う箇所も含まれているため、必ずしも人の判断が全て正しいということではないという裏付けとなった。トンネル延長100mの画像から壁面全面の危険箇所を検索する作業時間を1.0時間に対して、AIにて危険箇所を推定する場合の積み上げでは0.5時間に短縮され、生産性として200%の向上が見られた。

### 3) データプラットフォームへ危険箇所の連携データ出力技術

危険箇所をレーザー打音システムとデータ連携するため、危険位置の3次元解析を行い、特定様式で出力するようにした。画像計測システムに同時搭載したレーザー計測システムによって得られる3次元形状データを点群解析し、危険箇所位置を2次元から3次元座標に変換出力した。変換出力するデータは、レーザー打音システムで必要な危険箇所の角度や距離、高さ、およびトンネル内における相対位置データをデータプラットフォームに登録できるようにした。位置データとは別に、撮影展開画像データ、溶脱物位置データ、トンネル施工長データも登録するようにした。

### 4) 現場検証

山口県、静岡県、山梨県のフィールドにおいて、MIMM画像計測の事前スクリーニング結果から、AI危険箇所の判定技術を用いて実証試験を行った。

現状手法の劣化を画像から人力トレースする机上作業に比べ、AI支援により目視作業の代替とした場合約180%の生産性向上となった。事前スクリーニングを実施することで、正確な展開画像と変状位置の把握により、変状スケッチ作業の削減、新規変状箇所の調査時間の短縮、打音箇所の判断時間短縮に寄与する。

## (2) レーザー打音によるコンクリート内部欠陥の高精度判断に関する研究開発（量子科学技術研究開発機構）

### 1) レーザー打音検査に対応した健全度判定AIの開発

レーザー打音検査装置によって取得したコンクリート表面の振動データから、道路トンネル点検要領に準拠した多段階の健全度判定を行うAIを開発した。レーザー

打音検査装置により計測した振動データを時間分解フーリエ変換してスペクトログラム（時間・周波数・強度の3次元データ）を作成することでノイズと信号を判別し、かつスペクトログラムの特徴量をAIの学習及び健全度の判定に使用した。本AIに合わせて、AIの教育用データの収集、学習、および判定に用いるユーザーインターフェースをそれぞれ開発し、これらを統合することで、レーザー打音装置により熟練の検査員による健全度判定を再現した。

### 2) 自然欠陥を模擬したコンクリート供試体を用いた教師データの作成

コンクリートに発泡スチロール等を挿入することで作成する従来の単純な形状の模擬欠陥ではなく、コンクリートに掛かる圧力により発生させたリアルな欠陥を有し、かつ内部構造の明らかなコンクリート供試体を新たに開発した。図6に示すように、本供試体を熟練の打音検査員により3段階の健全度（Ⅰ：健全、Ⅱ：要観察、Ⅲ：修繕推奨）に分類し、それぞれにレーザー打音検査を行うことで、レーザー打音検査装置で取得した振動データを分類し、これを教師データとして、畳み込みニューラルネットワーク学習することで健全度判定AIの教育を行った（学習に用いた凡例は1,000程度）。

### 3) 現場検証

供用中のトンネルにおいてレーザー打音検査を実施し、健全度判定AIの判定精度の実証試験を行った。図7に示すように、取得した1,846点の振動データに対してAIでの判定を行った試験では、検査員による判定を基準として、過大判定7.3%、過小判定0.7%、ノイズ1.7%であり、AIによる検査員の判定結果の再現率90.4%を達成した。

### 4) 生産性の向上

従来のレーザー打音検査では、データを取得する時間に加えて、取得した振動データを全て手作業により再確認する時間を必要としている。例えば、10点×10点のレーザー打音検査を100回行った場合を想定すると、これまでのデータの取得と解析に掛かる時間の割合はおおよそ2:8である。判定にAIを導入することで、解析に掛かる時間は約半分になることが予測され、従来の作業時間全体を10時間とすると、AIの導入によりこれが6時間に短縮されるため、生産性は約170%に向上する。

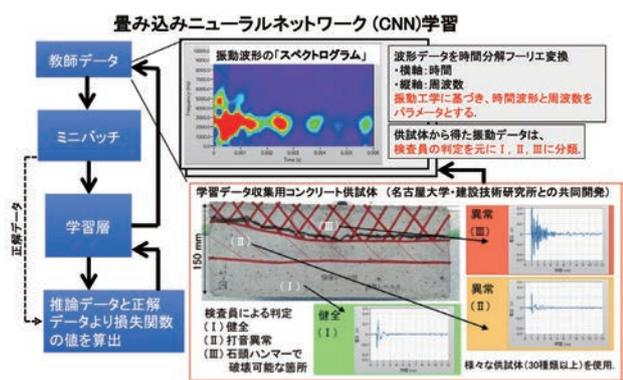


図6 自然欠陥を再現した供試体を用いた判定AIの学習



図8 画像データ (MIMM) と健全度判定データ (レーザー打音) の連携のためのデータベース

レーザー打音検査装置は、MIMMと連携することで打音検査を再現することを目的としており、それぞれのデバイスにAIを導入することで、総合的な生産性は現在の約190%に向上することが見込まれる。

### (3) インフラメンテナンスのためのデータベースの相互連携プラットフォームに関する研究開発 (社会基盤情報流通推進協議会)

#### 1) 画像計測データとレーザー打音データを紐づけるデータベースプロトタイプ構築

本データベースのプロトタイプ版は、①MIMM 走行計測後、分析・加工処理されたトンネル全体の撮影展開図 (画像データ) と危険箇所などの情報を取り込み、②トンネル全体やスパン別の危険箇所などの情報を検索、表示を可能とした。その結果、レーザー打音実施時に照射位置や範囲を効率的に特定させ、③レーザー打音結果も登録することにより、MIMMによる危険箇所とレーザー打音結果を重層表示できるシステムを構築できた (図8)。

#### 2) データベースプロトタイプ版の要件定義書等

上記のデータベース構築にあたり、MIMMなどによるトンネルの危険箇所に関する情報をデータベースとして取り込む際の、データ項目やデータフォーマットのデ

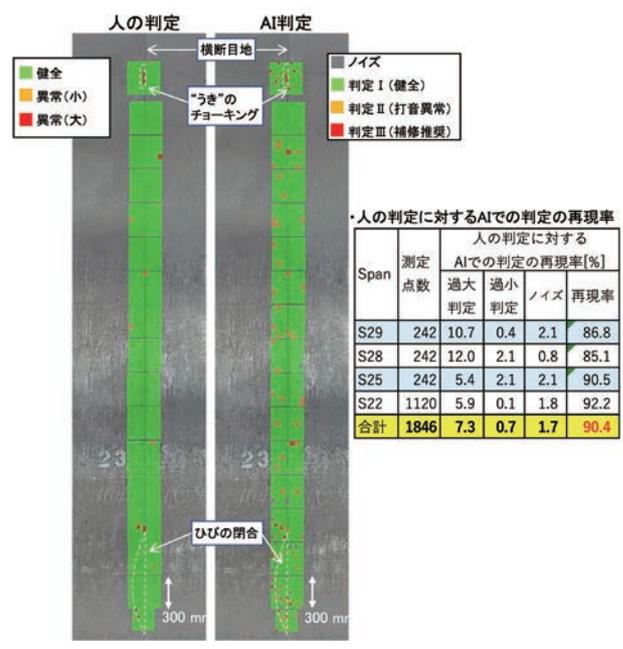


図7 健全度判定AIによるコンクリート健全度判定精度の検証

ータ定義を設定するなど、要件定義書等のドキュメントを作成した。

### (4) 工学的前兆判定AIを搭載したインフラ検査システム等の製品化と現場適用 (株式会社フotonラボ)

#### 1) 実証実験

研究開発段階での開発成果確認の実証実験は次のように行われた。

##### ①レーザー打音

2020年度は4件 (TRL4～5)の実証実験を行った (北海道:1件、九州:2件、静岡:1件)。2021年度については、3件 (TRL7)の実証実験を行った (高速道路:1件、一般国道:1件、地下鉄:1件)。2022年度は1件 (TRL7)の実証実験を行った (一般国道:1件)。

##### ②MIMM

2020年度は複数 (TRL4～5)の実証実験を、計測検査株式会社の受託業務に組み入れて現場にて実施した。2021年度も複数 (TRL7)の実証実験を、計測検査株式会社の受託業務に組み入れて現場にて実施した。

#### 2) 総合実証実験および公開

2022年度に研究段階の最終確認として以下のような

レーザー打音、MIMM、教育システムの総合実証実験を実施した。

#### ①プレ実証実験（2022年8月25日）

山口県矢地峠トンネルで技術連携の確認を行った。

#### ②総合的な公開実証実験（2022年9月6日）

静岡県焼津市の県道416号静岡焼津線の浜当日トンネルにおいて、関係省庁や地方自治体、建設コンサルタント、ゼネコンなど123名47機関を招いて、研究開発成果の総合的な発表と公開実証実験を行った（図9）。



図9 省庁・自治体・関連企業等を対象とした総合実証実験（静岡県焼津市浜当日トンネル）

#### ③フォロー実証実験（2022年10月18日）

山梨県広瀬トンネルにて、それまでの実証実験でやり残していた項目を最終的にフォローする実証実験を行った。本実証実験は、研究開発の原点となった、2012年12月2日に発生した中央高速道笹子トンネル事故（山梨県大月市、死者9名、史上最大の高速道事故）からの10年目の区切りとして、予防技術の進展を犠牲者の方々に報告する意味もあった。

#### ④情報発信

静岡での実証実験と山梨県でのフォロー実証実験の予告が読売新聞に大きく取り扱われた（2022年10月5日）。

山梨での実証実験はマスコミ各社（NHK・TBS・読売新聞・毎日新聞・山梨日日新聞など）に広く公開され大きく報道された。特にNHK総合テレビ「ニュースウォッチ9」では笹子トンネル事故10年特集（2022年12月1日放送）で、最新技術の開発成果として大きく取り上げられ、全国に広く情報発信された。

### 3) 海外事業調査企画

海外事業に向けては、「ODAにおける竣工検査に各種機械検査を導入する」、「現地作業員の指導に教育システ

ムを利用する」といった戦略を企画し、国家施策の指導下での展開を内閣府に報告した。

## 4 まとめと今後の展望

### (1) 画像計測による危険箇所の高精度抽出に関する研究開発（計測検査株式会社）

トンネル壁面の展開画像からAI抽出を施し、危険箇所を抽出するアルゴリズムとデータベース登録するまでの業務フローを構築した。今後もトンネル計測の実運用業務において、AIおよび危険箇所抽出を行う。データベース登録数の充実と、レーザー打音結果のフィードバックによる危険判定との因果関係を蓄積し、現場でのたたき作業時間短縮の精度向上に努める。また、今回現場検証において、トンネル壁面に描かれたチョーキング線の画像が不鮮明なことで抽出結果に影響することが分かったので、チョーキング線が一部欠如しているような場合でも予測抽出できるように技術改良を行うことで、より作業時間の向上が見込める。

### (2) レーザー打音によるコンクリート内部欠陥の高精度判断に関する研究開発（量子科学技術研究開発機構）

レーザー打音検査装置によって取得したコンクリート表面の振動データから道路トンネル点検要領に準拠した多段階の健全度判定を行うAIを開発した。供用中のトンネルに対する適用試験では検査員の判定結果の再現率90.4%を達成した。また、AIの導入により、MIMMとレーザー打音を組み合わせた打音作業の生産性は、現在の約190%への向上が見込まれる。今後は、さらなる精度向上のための学習データの蓄積を継続するとともに、本AIをレーザー打音検査装置に搭載し、計測時のその場診断にも活用する予定である。さらに健全度判定AIの教育用データの収集、学習、および判定UIを統合し、レーザー打音検査装置のオプションとして、ユーザーが現在運用している判定基準に則した判定AIを独自に構築するための統合UIを開発する予定である。

### (3) インフラメンテナンスのためのデータベースの相互連携プラットフォームに関する研究開発（社会基盤情報流通推進協議会）

MIMMで計測・処理した危険箇所データと、レーザ

一打音データを連携させるシステムを構築した。これにより、画像計測からレーザー打音の照射位置や範囲を効率的に特定ができた。今後は、実運用を重ねながら、本データベースの機能拡充や利用しやすさの向上に努めるほか、3次元データなどほかのデータベースとの連携についても引き続き検討していく予定である。

#### (4) 工学的前兆判定 AI を搭載したインフラ検査システム等の製品化と現場適用（株式会社フォトンラボ）

図4の全体戦略図をもとに説明する。

SIPの研究開発成果は社会実装責任組織である株式会社フォトンラボに各研究組織からすべて引き継がれ、製品化および各想定市場への営業が行われ、具体的な社会実装が実行される。フォトンラボは理化学研究所・量子科学技術研究機構の研究者と計測検査関係者の出資により設立され、半公的事業と認められ金融機関（埼玉りそな銀行）が指導的に直接出資し、事業戦略的に事業会社（JFEエンジニアリング、東京センチュリー）が出資して事業展開を支えている。

実際の市場への社会実装展開は、インフラの維持管理を入札する建設コンサルタントまたは鉄道・高速道路会社の系列保全会社を通して行われることがほとんどである。建設コンサルタント会社の大手である建設技術研究所とは業務提携を結び、ほかの大手会社（日本工営・応用地質・パシフィックコンサルタンツ）とも協力関係を築いている。

各市場別に社会実装展開の現状・方向性を説明する。

##### ①道路トンネル

国土交通省道路点検支援技術性能カタログ（トンネル）に掲載されており、一般公道での入札に採用可となっている。

MIMMとレーザー打音の連携による、AIを搭載した検査機器システムは最終年度でTRL7に達する見込みである。2023年度からは、積極的に計測サービスを有料で受託していく（TRL8）。本格的な普及は、2024年3月に想定される、国土交通省道路局が管轄する道路定期点検要領の改訂後と想定されている。

想定している顧客は建設コンサルタント各社、高速道路各社および系列保全会社、自治体である。

##### ②鉄道トンネル

道路での成果を水平展開する形で、開発テーマを国土交通省鉄道局と議論しており、具体的な開発は2023年

度から着手する予定である。すでに横浜市営地下鉄では現状の機器群を使って性能実験を開始している。

想定している顧客は、JR各社およびその系列保全会社、民間鉄道会社、地下鉄各社である。

##### ③導水路

民間企業の水力発電所の導水路改修に協力しながら、水中作業能力、超小型化など越えなければならない技術課題を検討している。この分野は、発電所の導水路にとどまらず、65万kmに及ぶ上水道、49万kmに及ぶ下水道という未着手の市場が存在する。想定している顧客は電力会社各社、自治体、水力発電所を持つ民間企業である。

将来的には自治体の上下水道が視野に入っている。

##### ④橋梁

国土交通省道路点検支援技術性能カタログ（橋梁）に掲載され、一般公道での入札に採用可となった。

道路トンネルへの社会実装の水平展開対象であり、最終年度には現場実証実験が可能となる（TRL6）。

本格的な普及は、2024年3月に想定される、国土交通省道路局が管轄する道路定期点検要領の改訂後と想定されている。

想定している顧客は建設コンサルタント各社、高速道路各社および系列保全会社、自治体である。

##### ⑤コンクリート構造物

岸壁の防波堤や建築物外壁、煙突などのコンクリート構造物で劣化診断ができるか否かの基礎実験が開始された（TRL6）。

##### ⑥自動運転のためのインフラ計測

インフラ維持管理データの自動運転用のマップデータへの反映を静岡県交通基盤部と情報交換を始めた（TRL3）。

##### ⑦環境計測 気象・汚染 防災

インフラ計測技術の環境及び防災への適用を静岡県交通基盤部と情報交換を始めた（TRL3）。

防災との連携が大きく期待され、静岡県浜当目トンネルでの公開実証実験時には、静岡県の防災点群データベース「バーチャル静岡」と大崩海岸付近でのトンネル維持管理データベースとのリンクテストデータが展示された。

##### ⑧海外戦略

ODAなどの国家施策に連動させて検討し、原則として日本の大企業との国内取引の形態をとる。特に、教育システムは海外市場が主であると考えられ、ODAと連動させた普及を目指す。

AIを活用したヒューマン・インタラクションによる取り組み

# Web 等に存在するビッグデータと応用分野特化型対話シナリオを用いたハイブリッド型マルチモーダル音声対話システムの研究

KDDI株式会社、国立研究開発法人情報通信研究機構、  
NECソリューションイノベータ株式会社

- 本研究開発では、日本において重大な社会課題の一つとなっている高齢者のケア・介護領域において、ケアマネジャー、介護作業員の作業負担軽減を担うと同時に、やはり社会問題として指摘されている、高齢者の社会からの孤立の回避に資するマルチモーダル音声対話システムの実現を目的とする。
- ケアマネジャーによる面談を対話システムで代替するために、高齢者の健康状態や生活習慣チェックを行う目的志向対話と、雑談を行う大規模データ駆動対話をハイブリッドに制御しながら介護モニタリングを遂行できる、音声対話システムを開発した。
- 実証により、目的志向対話によって必要な情報を高精度に取得できていることを確認。さらに、無味乾燥とならないように雑談対話を組み込み、多数のシステムの雑談的応答に対して、実際に笑顔や「(雑談中にシステムが話した生活のヒントなどに関して) 試してみます」のような、ポジティブな反応を高齢者から得られることを確認した。

## 1 研究の目的

本研究開発においては、高齢者介護分野における介護モニタリング\*1と呼ばれる面談作業を対話システムで一部代替し、同時に、健康状態悪化の原因となることが指摘されている、高齢者のコミュニケーション不足を抑制できるマルチモーダル音声対話システムを開発した。このシステムは、介護モニタリングにおける高齢者の健康状態、生活習慣チェックを行う対話に相当する目的志向対話と、コミュニケーション不足の抑制を目的としている。Web上の情報などを活用して行う大規模データ駆動対話（雑談）という、複数種の対話をハイブリッドに制御する介護モニタリングシステムとなっている。

開発にあたっては、介護の専門家の知見を積極的に対話システムに取り込んだ。介護職の業務負担を軽減させることで、深刻な社会問題になっている介護職の人手不足の解消や、健康悪化につながる高齢者のコミュニケーション不足、社会的孤立の回避に資するマルチモーダル音声対話システム「MICSUS (Multimodal Interactive

Care Support System)」の実現を目指した。

## 2 実施期間と方法

### (1) 実施期間

2018年度～2022年度

### (2) 実施方法

本プロジェクトは、図1に示す研究体制スキームを構築し、KDDI株式会社、国立研究開発法人情報通信研究機構(NICT)、およびNECソリューションイノベータ株式会社(NES)が共同で開発を行った。

本プロジェクトで構成する、MICSUSの全体構成図を図2に示す。左上の高齢者は、キャラクターAIなどを活用した高度対話UIデバイスや、スマートフォン、タブレットなどを利用した音声対話UIなどを介して、システムとマルチモーダル対話を行う。高齢者から得られたマルチモーダル情報は、マルチモーダル対話情報センシングモジュールなどによって音声認識、画像認識など

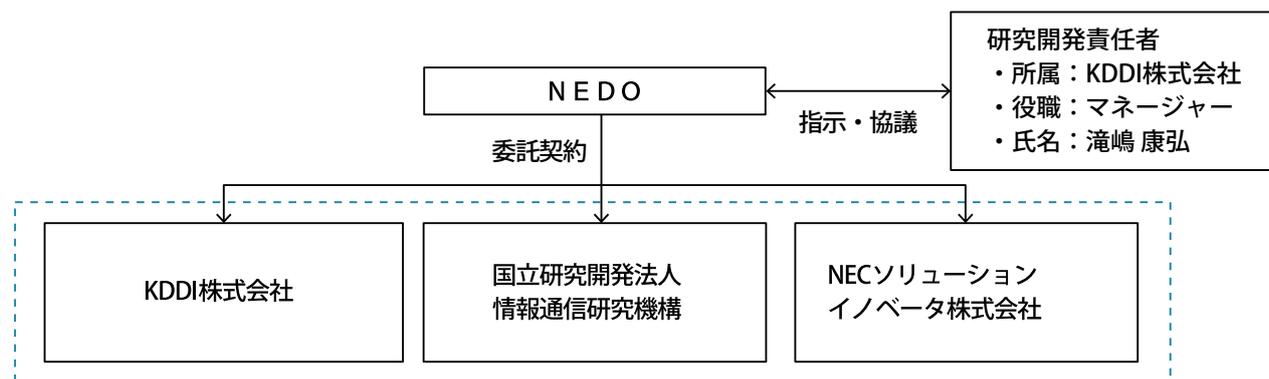


図1 研究体制スキーム

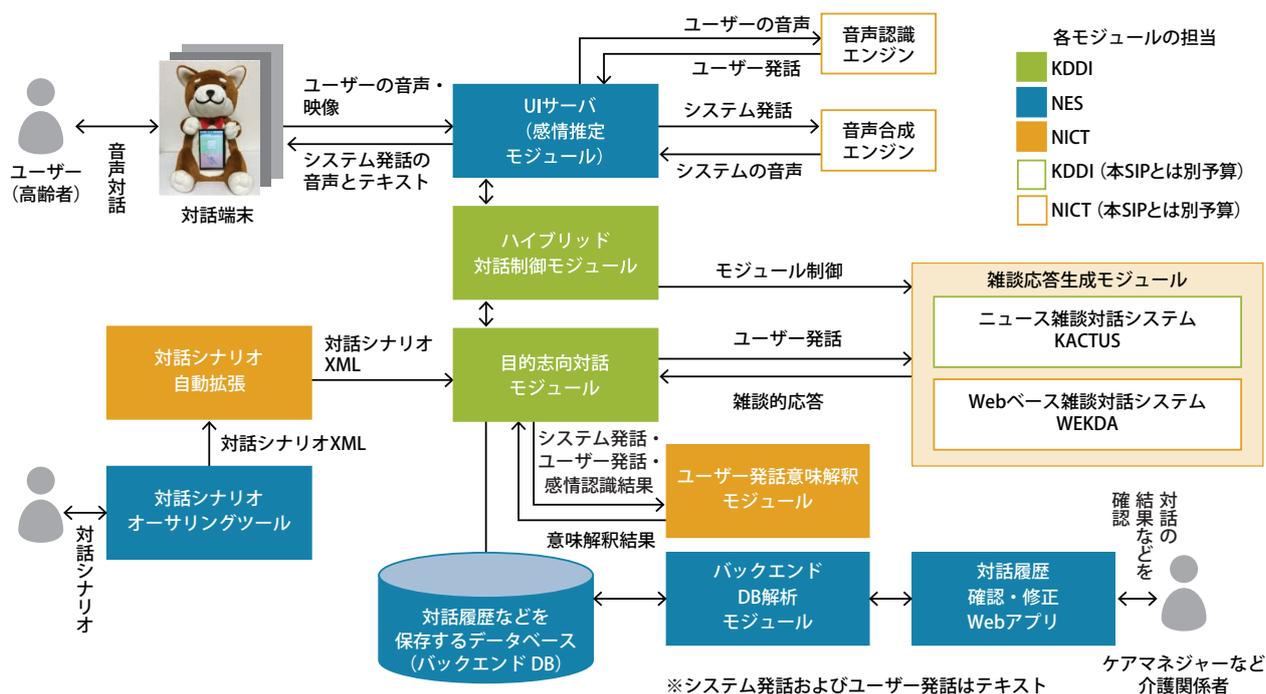


図2 本研究開発で構築する MICSUS の全体像

の認識処理が適用され、それらの認識結果は主としてユーザー発話意味解釈モジュールによって、柔軟かつ高度な意味的解釈を施され、目的志向対話モジュール、ならびに雑談応答生成モジュールに渡される。2つの対話モジュールは、それぞれ、対話シナリオ、Web情報などのビッグデータを用いてシステム応答の候補を生成し、状況に応じて、最終的な応答が決定される。

雑談応答生成モジュールはKDDIのKACTUS\*2、NICTのWEKDA\*3のそれぞれを拡張した2つのサブモ

ジュールからなっており、これらのモジュールもそれぞれ独立に応答を生成する。

目的志向対話モジュールが使用する介護用対話シナリオは、厚生労働省と株式会社日本総合研究所が策定・普及を進めている、「適切なケアマネジメント手法」に基づいて作成された。「適切なケアマネジメント手法」は、ハイパフォーマーな介護職のスキルを盛り込みつつ、高齢者の心身の状態に関して、疾患やそのステージごとに整理されたチェック項目を含むものである。具体的な対

\*1 ケアマネジャーが高齢者に直接会って、健康状態・生活習慣を面談でチェックする作業。

\*2 KDDIが開発した、ソーシャルメディア解析技術に基づく時事話題対話エンジンを活用し、利用者の興味関心に合致し社会で話題になっているニュースをリアルタイムに取り込む雑談対話のAIシステム。

\*3 国立研究開発法人情報通信研究機構（NICT）が開発した、利用者が自然な文を入力すると文章の意図を解析し、さまざまな話題やトピックに関して対話を展開できるシステム。

話シナリオ作成作業は、対話シナリオオーサリングツールを活用して人間が作成し、シナリオ自動拡張モジュールによって、さらに柔軟性を増大させた。

対話のログは、さまざまなデータと共にデータベース（バックエンド DB）に格納され、バックエンド DB 解析モジュールを介して、目的志向対話モジュールなどから対話シナリオによってアクセスされる。また、対話履歴確認・修正 Web アプリを介して、ケアマネジャーや高齢者の家族に伝達される。

ユーザー発話意味解釈モジュール、シナリオ自動拡張モジュールバックエンド DB 解析モジュール、雑談応答生成モジュールが必要とする学習データ、介護関連等用語辞書を含めた大規模言語資源を構築した。

実証実験については、高齢者介護などの応用分野に係る複数のユースケースについて、当該分野における事業者などの協力のもとで実施した。また、本研究開発の成果を広く活用していただくため開発者コミュニティを構築し、基盤となる技術の普及を図った。

以下に、主なサブテーマおよび実証実験の実施概要を示す。

#### 1) 高度マルチモーダル対話プラットフォーム

高齢者介護などの応用分野に特化した、対話シナリオを用いた「目的志向対話モジュール」と、目的志向対話モジュールでは対応が不可能な多様な入力に対しては、インターネット上のものも含めた多様な情報、知識を用いて、いわゆる「雑談」も含めた臨機応変な対応を行う「雑談応答生成モジュール」を組み合わせた、ハイブリッド型の MICSUS を実現するプラットフォームを開発した。

#### 2) 大規模言語資源の構築と深層学習を用いた高度対話・自然言語処理技術

本テーマでは、特に深層学習の活用が必須なユーザー発話意味解釈モジュール、シナリオ自動拡張モジュールを開発し、NICT の Web ベース雑談対話システム「WEKDA」をチューニングして、本プロジェクトの最終成果である MICSUS へと組み込む。また、これらの開発、チューニングにおいて必要な、学習データや辞書を含む大規模言語資源を構築した。

#### 3) 高度マルチモーダル対話インターフェース技術

MICSUS において活用するマルチモーダル情報を処理し、ユーザーの状態を推測する技術などを含む高度マル

チモーダル対話インターフェースを実現する。本テーマでは、対話システムのインターフェースとして想定される、マイク・カメラから入力される音声の特徴・映像（画像）によるユーザーの感情としぐさの推定技術、およびバックエンド DB に保存した対話ログをケアマネジャーや介護事業者が利用可能な技術を開発した。

#### 4) 介護用対話シナリオ構築技術

応用分野の現場において必要とされる、対話シナリオの設計と実装を行った。対話シナリオの実装のために、クラウド環境を利用した対話シナリオオーサリングツールを開発し、介護事業者によるシナリオの拡充を可能とした。そして、ケアマネジャーが自分の担当要介護者のモニタリング項目に応じて実行できる、シナリオセットを選択して提供した。また、対話シナリオについては、利用者の表情などのマルチモーダル情報に連動して進行できるものとした。

#### 5) 実証実験

実証実験により、「必要な情報の収集・記録に要するケアマネジャーの時間」の削減効果を検証した。具体的には、人のケアマネジャーによる情報の聞き取りに要している時間を、MICSUS で置き換えることによって時間を 10% 以上削減できることを確認した。ここで、「必要な情報」とは、「適切なケアマネジメント手法」にある情報項目と定義する。

#### 6) プロジェクト間協働の推進

本プロジェクト以外にも、SIP 第 2 期では介護分野におけるヒューマンインターフェース技術の開発などが進められている。SIP 全体として、それら他のプロジェクトの成果とのシナジーを極大化し、社会実装を早期に進められるようにするため、本プロジェクトと関連性の大きいほかのプロジェクトと協働し、「音声対話を用いたトイレ状況確認システム」「地域包括ケアシステムと MICSUS の連携システム」の開発と評価を実施した。

#### 7) 高齢者向け SNS と連携したスマートフォン上の高齢者向けマルチモーダル対話システム

スマートフォンで SNS にアクセスし、他者とコミュニケーションをとっているユーザーは数多いが、高齢者の多くはリテラシーの欠如などもあり、そうしたコミュニケーションを十分に行えているとは言い難い。本テ

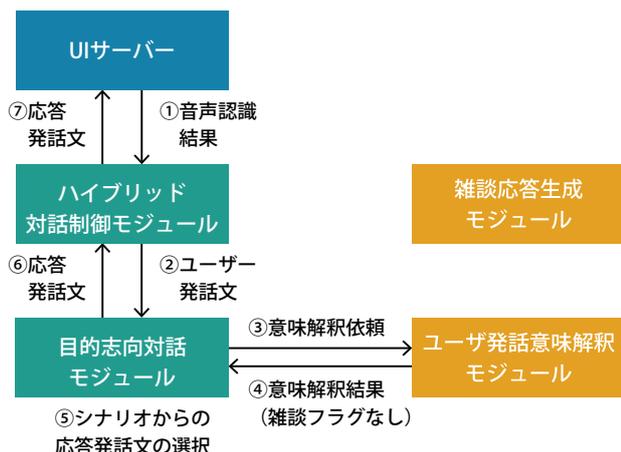


図3 シナリオに基づく発話生成フロー

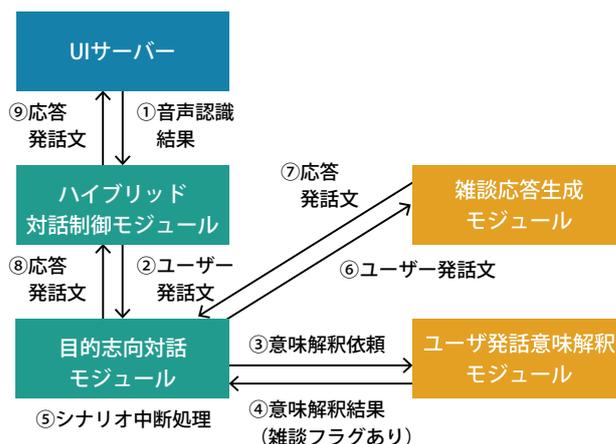


図4 雑談を含む発話生成フロー

マでは、高齢者が普段使っているスマートフォン上のMICSUSからSNSコンテンツを閲覧可能とし、対話で活用可能とすることでMICSUSとSNSを連携させ、健康状態悪化との相関があるかもしれない、高齢者のコミュニケーション不足を抑制するツールを開発する。具体的には、「MICSUSをベースに、高齢者と対話できるマルチモーダル音声対話インターフェースをスマートフォン上に開発する」、「そのインターフェースを介した対話・雑談によって、SNSの投稿やコメントなどのコンテンツを用いて対話・雑談の機会を増加させ、高齢者のコミュニケーション不足を解消すること」を目指した。

## 3 本研究の成果

### (1) 高度マルチモーダル対話プラットフォーム

マルチモーダル情報を入力として、シナリオに沿った対話や、ユーザーの発話内容に沿った雑談および話題提供のためのニュース提供を切り替えながら出力可能な、高度マルチモーダル対話プラットフォームについて説明する。

ここでは主要な構成要素となる、シナリオに従って対話を進行させる目的志向対話モジュールと、ユーザー発話の音声認識結果を逐次確認しながら、目的志向対話モジュールや雑談応答生成モジュールを連携動作させるハイブリッド対話制御モジュールの仕組みについて、それぞれ概要を述べる。2つのモジュールを中心とした処理フローを、図3と図4に示す。

### 1) 目的志向対話モジュールの仕組み

1回の対話における骨格となる、対話シナリオに基づいて対話を進めるモジュールである。XML形式で記述された対話シナリオ（詳細は「(4) 介護用対話シナリオ構築技術」に記載）を読み込み、条件分岐しながらシナリオを進めていくことが可能である。条件とは、ユーザー発話の音声認識結果以外に、感情推定モジュール（詳細は「(3) 高度マルチモーダル対話インターフェース技術」に記載）が推定したユーザーの感情や、カメラ画像から認識したジェスチャー情報、および過去の対話でユーザーが発話した内容、既往症の有無などの前提条件を参照できる。条件によって、シナリオの実行有無の変更も可能であり、例えば、ポジティブな感情が推定されると、関連する質問を追加することや、ネガティブな感情が連続して発生した場合にシナリオを途中で止めることもできる。

ユーザーの発話内容の解釈については、ハイブリッド対話制御モジュールが出力した、文単位の音声認識結果をユーザー発話意味解釈モジュール（詳細は「(2) 大規模言語資源の構築と深層学習を用いた高度対話・自然言語処理技術」に記載）に入力し解釈させる。解釈結果に応じて、相槌生成機能によって生成した相槌の文言を出力することや、次の質問へ移行する際のつなぎの文言をシナリオに従って決定している。

このとき、ユーザー発話意味解釈モジュールからの応答に、雑談することを判断したことを示すフラグ（雑談フラグ）が含まれていた場合、雑談応答生成モジュールにユーザー発話の音声認識結果を送信し、雑談応答生成モジュールから出力応答文を受け取る。出力応答文を受

け取ると、シナリオ対話を一時停止状態として、どこまで対話が進んでいたかを記録しておき、雑談応答生成モジュールの出力応答文をハイブリッド対話制御モジュールに送信して発話させる。さらに、ユーザーからの発話が続いた場合にも、上記と同様にユーザー発話意味解釈モジュールに入力し、雑談フラグの有無で雑談応答生成モジュールを継続して呼び出すかを判断する。雑談フラグが無くなった場合、記録していたシナリオ対話の中断地点に戻り、「それでは私からの質問に戻りますね」などのつなぎの文言を挿入した上で、続きのシナリオを実行する動作に戻る。

### 2) ハイブリッド対話制御モジュールの仕組み

逐次入力される音声認識結果を文章単位にまとめ、さらに複数の文章をまとめてユーザー発話として生成し、それを目的志向対話モジュールに出力する。例えば、『「そう」「だね」(短い無音区間)「朝ご飯」「を」「食べたよ」(長い無音区間)』といった音声認識結果を、「そうだね。」「朝ご飯を食べたよ。」の2文にまとめ、無音区間の長さに基づいてこの2文を1文にまとめるかどうか判断する。

また、端末のマイクのオン/オフのタイミングと目的志向対話モジュールの応答生成タイミングを決定する機能を持つ。マイクオフのタイミング制御により、システム発話がマイクに流れ込むことを防止し、音声認識の誤りを防止している。応答生成タイミングについては、無音区間の長さを閾値で判断することによって決定している。

加えて、ユーザーとの対話開始時に、ハイブリッド対話制御モジュールがバックエンドDB(詳細は「(3) 高度マルチモーダル対話インターフェース技術」に記載)から、過去の対話履歴やユーザーの趣味嗜好情報を取得し、雑談応答生成モジュールに提供している。

### 3) 利用者発話後からシステム応答までの待ち時間短縮に向けた取り組み

利用者ごとに対話のペースが異なることから、応答までの待ち時間は、人により短くする、長くするなど調整可能となることが望ましい。そのための取り組みとして、調整幅を広げるための応答時間短縮に向けた、投機的実行機能について検証した。

投機的実行機能とは、音声認識機能による利用者の発話終了の判定を待たずに、発話途中の情報を使って意味解釈などを前もって実行する機能である。これによって、

発話終了の判定と同時に、応答が生成できている可能性が期待できる。実証の対話ログ4名の46発話で検証し、ユーザー発話完了時点からシステムが応答生成するために2.5秒程度かかるようなケースでも、平均0.68秒程度に短縮可能な見込みが得られた。

一方、システム全体の処理負荷が増えることから、要求される応答速度やサーバースペックに応じて投機的実行機能の実装可否を調整することが望ましい。

### 4) 最新の話題の提供機能

ニュース記事から最新の話題を含んだ対話を提供するモジュール(KACTUS)をMICSUSに接続し、対話の最後に最新話題を紹介できるようにシナリオを構築した。高齢者3名に対する28件の最新話題の提供に対する実証実験の映像ログで検証し、約74.1%の割合でポジティブな表情になっていることを確認した。

また、最新話題を提供された高齢者が「参考になります」など、興味を持つような応答を返した割合が約63.0%であった。この結果から、最新話題の提供によって、例えば継続意欲につながるなどの効果を期待できる結果となった。

## (2) 大規模言語資源の構築と深層学習を用いた高度対話・自然言語処理技術

ユーザー発話意味解釈モジュールやシナリオ自動拡張モジュール、Web上のテキストを用いた雑談対話に関する成果を述べる。

MICSUSは「適切なケアマネジメント手法」に基づいて作成した質問を行い、ユーザーは音声発話で質問に回答する。本プロジェクトで開発したユーザー発話意味解釈モジュールは、ユーザー発話を解釈するモジュールである。本モジュールは、MICSUSの質問文と、NICTの音声認識エンジンによって認識されたユーザー発話文を入力として、以下の5種類の解釈結果を出力する。

- YesやNoなどで回答可能な質問である「Yes / No 質問 (例: 毎日三食ご飯食べていますか?)」に対する、ユーザー発話の分類結果。「Yes」、「No」、「不明 (回答が分からない)」、「前提矛盾 (質問の前提を否定する回答)」、「回答拒否」、「その他」の6種類に分類される。
- 名詞や短い文で回答可能な質問である自由回答質問 (例: 昨日の夕飯は何を食べましたか?) について、回答の有無の判定結果と回答キーワードの抽出結果。

- 雑談を展開すべきかの判定結果。
- 回答の訂正情報の認識結果。
- 後で聞く予定だった質問の回答をユーザーが先に発話した場合に、それを認識した結果。

これらの判定や抽出は、巨大ニューラルネットワークBERTをベースとして、NICTで独自に開発した音声認識誤りに頑健なニューラルネットワークであるHBERT（Web上のテキスト約350GBで事前学習）と、専用に構築した累計250万件以上の言語資源によって実現した（プロジェクト終了時点で言語資源の総数は累計300万件の予定）。

HBERTは、BERTを拡張したものであり、通常の入力である仮名漢字交じり文に加えて、読みの情報を音韻的な情報として入力することが特徴である。HBERTによって、「お水をよく飲んでいきますか？」に対して「飲んでます」と回答したが、音声認識誤りによって音が似ている「飛んでます」と認識された場合でも、正しくYesと解釈できる。また、大規模なWebテキストで事前学習したHBERTと、大規模な学習データによって「毎日三食ご飯食べていますか？」に対して、「胃腸の調子がいいからね」のように遠回しな言い方で回答しても、正しくYesと解釈できる。なお、さまざまな意味解釈を行うため、本モジュールは17種類のサブタスクを同時に実行でき、また一度の解釈に要する時間はミリ秒単位と非常に軽量である（図5）。

次に、シナリオ自動拡張モジュールについて説明する。対話シナリオは、「適切なケアマネジメント手法」に基

づいて人手で整備するが、人手による作業だけで柔軟な対話を実現することは難しい。例えば、ユーザー発話の解釈がうまくできなかった場合には、全ての質問に対して再度同じ質問を聞く「回答の聞き直し」というやりとりが発生し得るが、そのやりとりに対応する対話シナリオを人手で逐一追加するのは高コストである。そこで、本モジュールによって人が書いた対話シナリオを自動的に拡張し、回答の聞き直しのように対話中で広範に起こり得る一般的なやりとりを、人手による記述作業なしで実行できるようにした。このシナリオ自動拡張モジュールによって、対話中の任意の箇所が発生し得る、雑談エンジンによる雑談対話、回答を訂正する対話などを行う対話シナリオを自動的に追加することで、柔軟な対話を実現した。

2022年度は高齢者13名を対象として15日間毎日MICSUSと対話する長期実証と、高齢者36名を対象として1日に4回対話を行う短期実証を行った（合計で高齢者は49名、対話は344セッション、対話時間は合計約35時間、MICSUSと高齢者の対話ターンの合計は10,192回〔うち高齢者は4,924回〕、高齢者のターンはセッションあたり約14回。今後も被験者数を増やし、年度末までに約100名の高齢者による評価となる予定）。総合的な評価として、高齢者の発話に対して適切な応答が出力できたかを評価した結果、高齢者のターン数4,924回に対して、雑談的応答なども含めて93.2%の精度で適切な応答を出力することができた。評価は3名の評価者による多数決で決定した。Yes /

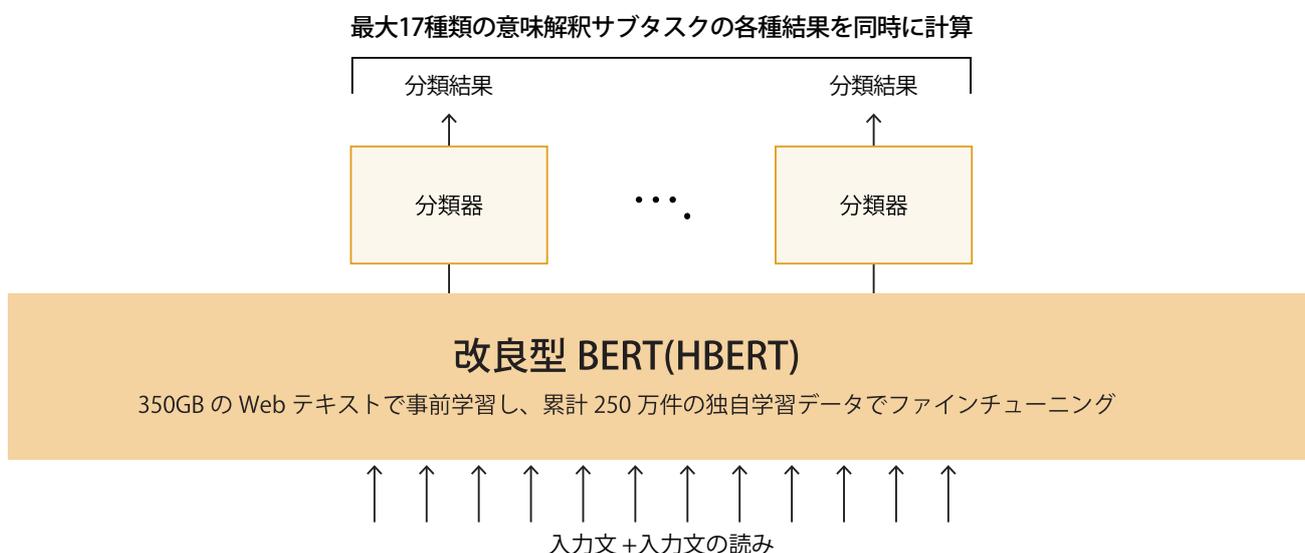


図5 ユーザー発話意味解釈モジュールで活用されているHBERT

No 質問については、高齢者から得られた回答を3名で評価し多数決をとったところ、正解率は94.0%であった。誤りの原因について、音声認識誤りの問題は前述のHBERTによって大幅に軽減されているが、HBERTでも解決できない著しい音声認識誤りが発生する場合もあり、これが不正解の原因と考えられるケースが最も多かった。実証実験では後述の通り、ケアマネジャーがYes/Noの解釈結果等を確認し、必要があれば修正を行ったが、修正が必要な箇所はごく僅かで、実用上問題のない精度が実現できることが分かった。自由回答質問については、ユーザーが質問に回答したかどうかの判定の正解率は93.7%、回答キーワードの抽出の正解率は83.2%であった。回答キーワード抽出の不正解の主な原因は、ユーザー発話中の回答キーワードなど重要な部分の音声認識誤りによるものが最多であった。

MICSUSは、ユーザー発話意味解釈モジュールによって、入力されたユーザー発話を雑談に展開すべきか否かが判定でき、Web上の情報をもとに雑談的応答を生成する音声対話システムWEKDAと連携することで、雑談ができる。前述の高齢者49名に対する実証実験では、合計451件の雑談的応答を生成できた。3名の評価者が対話の動画を見つつ、ユーザーの反応や雑談的応答の質を独立に評価し、多数決で評価を確定させたところ、90.9%が雑談的応答として適切であると分かった。さらに、ユーザーが笑顔を見せる、もしくは「やってみます」、「それは面白いですね」など積極的に興味を示す応答をするなど、積極的に興味を示したという厳しい条件での判定では、54.3%の雑談的応答がこれらの条件を満たしていた（笑顔は全体の31.0%、積極的に興味を示す応答をしたのは全体の40.0%）。MICSUSの雑談的応答は、既存の対話システムによく見られる単なる相槌的応答であったり、おうむ返しではなく、ユーザーの発話に関するチャンスやリスクなど、Webから取得された何らかの新規情報を含む設計になっているが、そうした雑談的応答としては良好な品質であることが分かった。加えて高齢者32名に対する短期実証では、雑談機能を有効にした場合の対話と、無効にした場合の対話（健康状態チェックのみが行われる対話）の順序を入れ替えて、それぞれ2回ずつ実施してもらったABテストを行い、どちらが良いかを5段階（スコア4、5が雑談機能がある方が望ましい）で評価するアンケートを行った。その結果、5%の有意水準で雑談機能がある方が良い（ $p$ 値=2.5%）という結果が得られた。スコアの分布は、スコア1:0名、

スコア2:1名、スコア3:9名、スコア4:9名、スコア5:13名で、スコア5が最多であった。雑談機能がある方が望ましいと回答した高齢者は69%であった。

### (3) 高度マルチモーダル対話インターフェース技術

本プロジェクトにおいて開発した、感情やしぐさの推定技術は、音声対話端末のカメラとマイクを用いて収集したユーザーの会話に関するマルチモーダル情報をリアルタイムに分析し、対話の応答で活用する。またバックエンドDBとその分析技術は、保存した対話履歴を分析し、ケアマネジャーなどによる対話履歴の容易な閲覧、確認、修正に活用する。

#### 1) 表情からの感情推定技術

近年の表情感情推定AI（既製品）ではFACS（Facial Action Coding System）理論に従い、表情筋の動作AU（Action Unit）をもとに深層学習のモデルによって感情を推定している。例えば、喜びの感情は、笑顔の表情筋の動作（頬や口角が上がり、口が開く）を深層学習し、推定する。

FACS理論は一般的な表情筋の動作と感情の関係によって成り立っており、加齢に伴う表情筋の基本位置の差異など、個々のユーザーに特有な表情の特徴を考慮していない。例えば、初見では怒っているように見える人が、家族など見慣れた人には普段通りに見え、平常とみなすといった事態があり得る。そこで、表情感情推定AI（既製品）が推定した感情を、ユーザー個人のデフォルトな表情に照らして、より適切な感情の認識結果に補正する個人バイアス補正技術を開発した（図6）。

対話中に、ユーザー個人のFACS理論によって認識した表情から推定される感情の認識結果を、個々のユーザーに頻出する表情から推定される感情の認識結果を用いて補正した結果、適合率91.2%、再現率93.4%を達成した。

#### 2) 音声的特徴からの感情推定技術

音声的特徴感情推定AI（既製品）は、機械学習などを用いてユーザーの音声の抑揚やピッチなどから感情を推定している。そこで本プロジェクトでは音声対話端末のマイクをシステム発話時はOFFにし、ユーザーの音声以外を収集しないようにした。さらには、収集した音声を発話検知VAD（Voice Activity Detection）を用いて

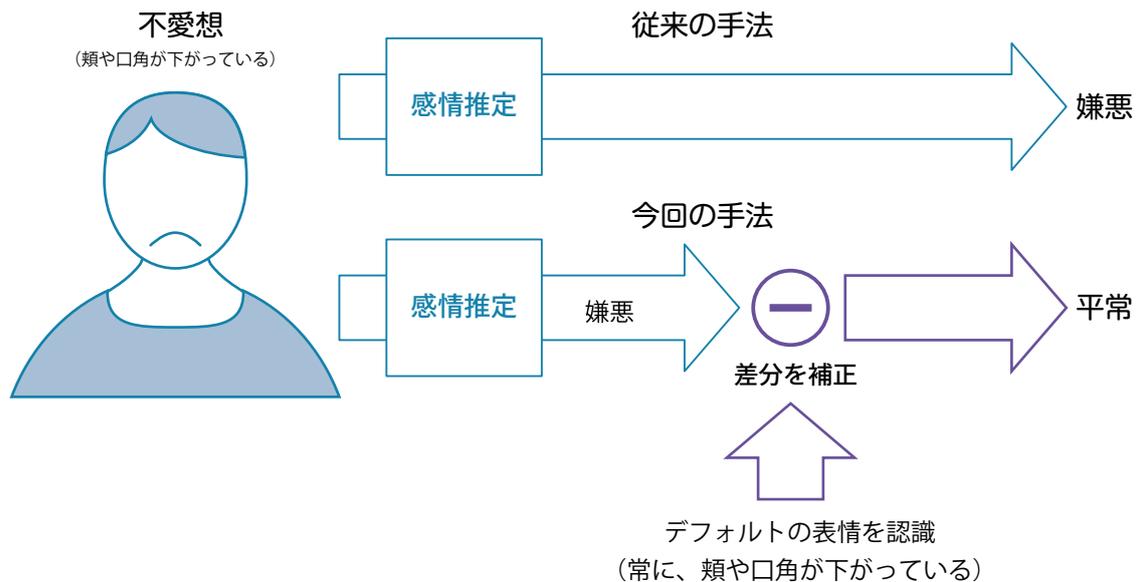


図6 個人のデフォルトな表情を用いた感情推定のバイアス補正技術

ストリーム形式の音声データからリアルタイムに発話区間を切り出して感情推定した結果、適合率 87.0%、再現率 87.0%を達成した。

### 3) 統合感情推定技術

人は表情や音声的特徴など複数の非言語情報を一つにまとめて対話に利用していることから、前述の独立したシングルモーダルから推定した「表情感情」と「音声的特徴感情」をさらに一つにまとめる必要がある。そこで、同じ発話区間で推定した感情の推定確度、ならびに表情の変化量や音声の音量などを考慮し、表情か音声のどちらかの感情を対話に用いるかを決定した。また、対話においてユーザー発話にシステムが応答するまで間延びしないよう、表情や音声的特徴による感情推定を並列処理するとともに、統合感情推定処理の高速化することで、ユーザー発話後に平均 15.4 ミリ秒で統合感情の結果を得ることができた。

### 4) しぐさ推定技術

ユーザーの頭部の傾きと顔の特徴点を利用し、しぐさ(うなずき、首振り)を認識する技術を開発し、システムの質問に対してユーザーが肯定的か否定的かの判断に大きく明らかなうなずき・首振りを活用した。大きく明らかなうなずき・首振りの適合率と再現率は 100%を達成した。

### 5) バックエンドDB および分析技術

ユーザーの発話内容や意味、感情ならびにシステムの応答文や対話シナリオの進捗情報をイベントドリブンで保存し、保存した情報を同じ事象などと関連付けて取り出せるデータベースを構築した。また音声対話装置がリアルタイムに収集する、ユーザーの動画や音声も、システムが分析した対話情報と関連付けて保存している。これら保存した対話の情報を、ケアマネジャーなどが容易に利活用するために、以下を可能にした。

- ユーザー発話の内容や意味、感情と、発話時の動画の関連付け。
- 質問と意味解釈の結果を対にした要約。
- ケアマネジャーなどによる意味解釈結果の手動修正。
- 会話全体のレポート出力。

さらに、情報利用するユーザーごとにアクセス権を設定でき、会話情報の保護とマルチテナントの実現を図った。

## (4) 介護用対話シナリオ構築技術

### 1) 対話シナリオ作成

システムのユーザーである要介護高齢者が受け入れやすく、かつケアマネジャーなどの専門職が必要とする情報を収集できるような対話シナリオを作成した。具体的には、以下に示す 3 ステップで実践的なシナリオを作成した。

第 1 ステップとして、シナリオの構造の検討および生成するシナリオの検証のため、一般社団法人日本介護支援専門員協会および株式会社インターネットインフィ

ニティの協力を得て、人の専門職が実施する在宅介護モニタリング対話計 90 件（1 件約 30 分）の音声・動画データを収集した。それらのデータに対して、モニタリング対話 1 件当たり 3～5 問の主要な質問を担当ケアマネジャーが予め想定して臨んでいること、ケアマネジャーが聞きたい質問の聞き取りに入る前に導入（アイスブレイク）を実施していること、一つの主要質問を聞き取りするために、関連する質問や雑談を挟んだ聞き取りをしているといった構造を確認した。

第 2 ステップとして、次項で開発した対話シナリオ記述言語仕様をもとに、「適切なケアマネジメント手法」に基づいて約 6,000 件の質問（ケアマネジャーが要介護高齢者に聞き取りする主要な質問）のシナリオを作成した。具体的には、一つの質問に対して「平常」、「不快」、「不快が直近に連続」、「不快が同じ質問で複数回」、「喜び」、「驚き」の 6 状態のシステム応答をそれぞれ定義した。ここで作成した約 6,000 件の質問は、ユーザー実証の結果を通じて都度ブラッシュアップを重ねるとともに、新型コロナウイルス感染症のまん延を受けて国立研究開発法人国立長寿医療研究センターが公表した、高齢者向けの新型コロナウイルス感染対策ガイドラインを

踏まえた質問項目の追加（12 問）も実施した。

第 3 ステップとして、モニタリング対話 1 件当たり 3～5 問の主要な質問をケアマネジャーが事前に想定していることを踏まえ、約 6,000 件の質問の中から 1 件当たり 3～5 問の主要な質問を中核に対話の流れを組み立てた「プリセット」を 15 件作成した。ここで作成したプリセットは、ユーザー実証（15 日間の連続利用）で利用し、本研究全体で開発を進めるシステム全体の評価に活用した。その実証に先立ち、プリセットで選択した主要な質問（プリセット 1 件につき 3～5 問）の組み合わせ方や質問のつなぎ方について、業務経験年数の長いケアマネジャー 8 名が評価し、延べ 60 回以上の試行実施を経てプリセットシナリオのブラッシュアップを実施した。なお、プリセットを複数作成するにあたっては、後述する「3）対話シナリオオーサリングツール開発」に示すオーサリングツールを活用し、シナリオ製作者の作業時間が大幅に短縮されることを確認した。

## 2) 対話シナリオ記述言語仕様策定

マルチモーダルを考慮した対話を可能とする、XML を用いた対話シナリオ記述言語仕様を策定した。Yes や

シナリオ編集ツール Ver10.19 - マスターシナリオ編集

ファイル: モード選択, 保存

メニュー: 検索, 設定更新

ユニット一覧:
 

- 基本ケア
- 開始終了
- 個人属性
- 基本情報(家族)
- 導入

質問の基本設定:
 

- 優先度: 10
- ID: 基本-1-1-2
- モニタリング項目: 本人・家族の日常の食事に対する意識や食生活の内容
- モニタリングで確認すること: 食事の回数 (3食食べているか)
- 質問間隔: 毎日
- 質問の前提条件: 特になし
- クッション発話(大): 普段の食事についてお聞きします。
- クッション発話(小): 次に進みます。

ページ: エレメントID, スタートID, 呼び出し元

戻る: 基本-1-1-2-現状, 基本-1-1-2-現状-MS

質問	ユーザー回答	感情	Positive	発話条件	システム発話 (応答)	モニタリング結果	遷移先
毎日しっかり食事をしていますか？	Yes(はい)	平常	1	特になし	毎日しっかり食事をしているんですね。	1日3食食べている	ElementEND:
			0	特になし	毎日しっかり食事をしているなら、安心ですね。食事の時間が楽しみです。	1日3食食べている	ElementEND:
			0	特になし	そうすると、以前よりも毎日しっかり食事をするようになったんですね。	1日3食食べている	ElementEND:
		0	不快	特になし	それは、以前よりも毎日しっかり食事を摂られて、体もちついたらそうですね。	1日3食食べている	ElementEND:
		0	不快が直近に連続	特になし	お気を悪くされたいしません。食事はしっかりしてらんですね。	1日3食食べている	SessionEND:Unit:開始終了-終了-最終-2
		0	不快が同じ質問で複数回	特になし	何度も同じ質問をして、すみません。ちょっと分かりづらかったですかね。	1日3食食べている	SessionEND:Unit:開始終了-終了-最終-2
	No(いいえ)	平常	1	特になし	毎日しっかり食事をしてらんですね。それに今日はご機嫌良さそうですね。	1日3食食べている	ElementEND:
			1	特になし	急に聞いてごめんない。毎日しっかり食事をしているのなら、安心ですね。	1日3食食べている	ElementEND:
			0	特になし	毎日しっかり食事をしてないんですね。	1日3食食べていない	ElementEND:
		0	特になし	毎日しっかり食事をしないと、栄養が足りないかもしれませんね。	1日3食食べていない	ElementEND:	
		0	特になし	毎日しっかり食事をしてないんですね。	1日3食食べていない	ElementEND:	
		0	特になし	毎日しっかり食事をしないと、栄養が足りないかもしれませんね。	1日3食食べていない	ElementEND:	

図 7 シナリオ編集ツール

Noなどで回答可能な質問の回答を意味解釈した結果となる、「Yes」、「No」、「不明」の3状態それぞれにシステムからの応答が記述可能であることに加え、「質問への回答時における高度マルチモーダル対話インターフェース技術」で説明された手法で推定した、4種の感情（平常、不快、喜び、驚き）を用いた。さらに、高齢者の継続利用を促すために、対話に用いる感情として「平常」、「不快」、「不快が直近に連続」、「不快が同じ質問で複数回」、「喜び」、「驚き」の6状態それぞれに、システムからの応答が記述可能である。質問をスムーズに行うために、対話のクッションとなる発話をシナリオに記述でき、この発話も4種の感情（平常、不快、喜び、驚き）によって切り替え可能である。このような、回答の意味3状態と感情6状態を組み合わせ、次に説明する対話シナリオへの遷移についても、対話シナリオ記述言語仕様にて定義した。

なお、始まりの挨拶、終わりの挨拶を記述する開始・終了用シナリオ記述言語仕様、および個人の趣味趣向をヒアリングする個人属性シナリオ記述言語仕様も策定し、質問のみでなく対話の中でユーザーの情報を収集して雑談などの対話を可能とした。

また、介護質問の前にユーザーを和ませるアイスブレイクを行う導入シナリオを記述でき、前回のポジティブな回答や前回の対話の概要を話題とした対話を可能とした。

### 3)対話シナリオオーサリングツール開発

対話シナリオ記述言語仕様に沿った対話を行うための質問文や、質問の遷移を容易に作成可能とするオーサリングツール（Windows アプリ）を開発した。オーサリングツールは、シナリオ編集ツールとシナリオ作成ツールがあり、作成したシナリオはクラウド環境に保存する。

シナリオ編集ツールでは、エディタ画面にて質問文と応答文を入力でき、応答文は回答の意味3状態と感情6状態の組み合わせごとに記述できる（図7）。また、質問文と応答文の組み合わせをユニットという単位で管理し、その中で「クッションとなる発話」、「質問文」、「質問への応答文」、「モニタリング確認」、「次に実行する質問を指定」を定義できる。さらに、質問と応答を音声で確認できる。

シナリオ作成ツールでは、シナリオ編集ツールで作成したユニットをつなぎ合わせて対話フローを設定でき、対話フローは回答の意味3状態と4種の感情の組み合

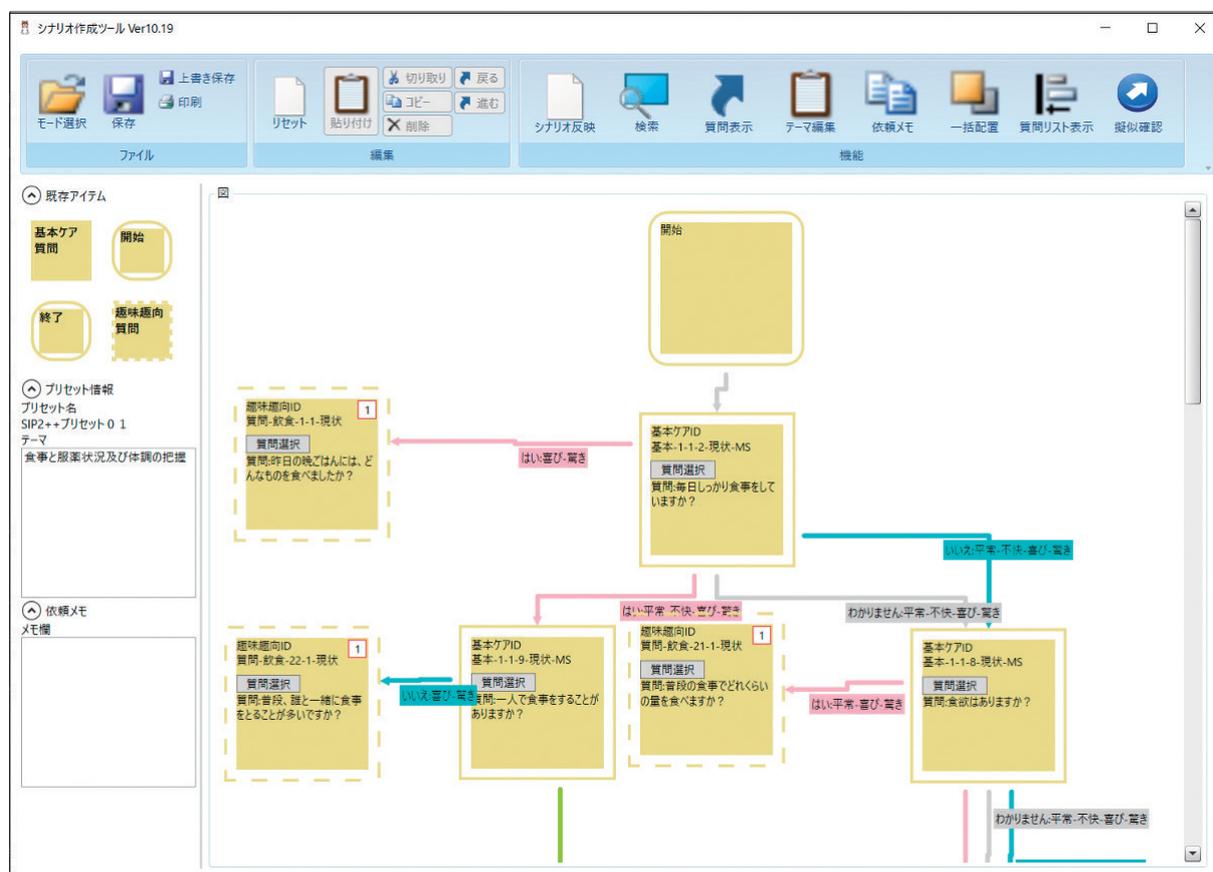


図8 シナリオ作成ツール

わせにて分岐可能である（図 8）。

質問となるユニットをアイテム選択画面から図にドラッグ&ドロップし、実行する質問を選択。そして、それらのユニット間の遷移条件を「Yes」、「No」、「不明」、および「平常」、「不快」、「喜び」、「驚き」で選択可能として、質問の回答と感情により一連の対話を対話の実行単位であるシナリオセットとして構築可能とした。

開発した対話シナリオオーサリングツールを利用した対話シナリオのカスタマイズをケアマネジャーに実施してもらい、シナリオ作成に対話シナリオオーサリングツールが有用であることを確認した。

## (5) 実証実験

本研究で開発する対話 AI の有効性を検証するため、要介護高齢者を対象としたユーザー実証、およびケアマネジャーを対象としたユーザー実証を段階的に企画、実施した。研究開発の段階に応じて、小規模なユーザー実証を並行的に実施することで研究開発の効果を高めるとともに、新型コロナウイルス感染症対応を挟んで、2020 年度から 2022 年度にかけて要介護高齢者が集住する施設（サービス付き高齢者向け住宅、ケアハウス）の入居者、および在宅で生活する要介護高齢者を対象として、15 日間毎日 1 回、連続して利用してもらう実証を 23 回実施した。連続実証では、15 回異なるプリセットシナリオを実行し、要介護高齢者単独でシステムとの対話を実行してもらった。対話結果は担当のケアマネジャーに確認してもらい、「適切なケアマネジメント手法」に基づいた要介護高齢者の幅広い状況を把握できること（ケアマネジャーが求める情報を収集できていること）を検証した。

本来、MICSUS と同じ情報を人間が対面で聞き取る場合の所要時間とその記録時間を測定し、それらの時間と、MICSUS を活用した Web アプリでの対話履歴確認や修正に要した時間との比較では、平均 7.0 分から 2.2 分（▲ 4.8 分）に短縮できることを確認できている。これは、面談 1 回に相当する情報収集におけるケアマネジャーの業務時間を、およそ 69% 効率化できる計算となる。上記は、ケアマネジャーの訪問回数を減らさずに MICSUS を活用する保守的な想定だが、訪問回数を削減すれば「情報収集以外の話題を話す時間」に加えて往復の移動時間も含め、さらに多くの業務時間が削減可能と考えられる。

また、要介護高齢者を対象としたユーザー実証に加

え、さらに活用が見込まれる産業分野をデスクトップ調査し、対話技術が活用される可能性のある、高齢者介護以外の応用分野の候補を抽出した。具体的には、要介護状態の手前段階である、フレイル状態にある高齢者を対象としたユーザー実証を実施した。要介護高齢者向けの 15 回連続の実証とは別に、2 回程度のプリセットを実施し、対話システムを通じて健康状態や生活状況などを把握した。これによって、健康状態や生活状況のモニタリング場、高齢者の集合住宅での入居者の健康管理、自治体の地域包括支援センター、および保健所における介護予防事業のフォローアップなどに利用できる可能性を検証した。

## (6) プロジェクト間協働の推進

対話システムの応用範囲の拡張を目指し、関連プロジェクトのユースケースに対話システムを適応させ、有用性を検証する活動を 2 件実施した。

### 1) 音声対話を用いたトイレ状況確認システム

株式会社 aba と連携し、介護施設などの共用トイレを利用する高齢者を MICSUS が声でサポートするユースケースに対応したシステムを構築して、対話システムの有用性を検証した。着座センサーなどの IoT センサーと連携するシステムを構築し、実際の介護施設のトイレで動作検証を実施した。着座をきっかけとして、「紙はありますか？」などのよくある困りごとについて、システムから能動的に話しかけることを可能にした。

### 2) 地域包括ケアシステムと MICSUS の連携システムの開発と評価

株式会社アルムの地域包括ケアシステムと連携し、ユーザーの MICSUS による対話終了後に、対話履歴を地域包括ケアシステムへ自動登録できる機能を開発し、地域包括ケアシステムがリスク分析に必要とする質問項目に関する回答データ、ならびにユーザー発話を再確認できるよう対話時の動画を地域包括ケアシステムへ登録可能にした。

## (7) 高齢者向け SNS と連携したスマートフォン上の高齢者向けマルチモーダル対話システム

### 1) MICSUS サーバーと連携するマルチモーダル音声対話アプリの構築

MICSUS の汎用性を高めるための施策として、市販の

スマートフォンをユーザーインターフェースとして活用するための開発を、FCNT 株式会社と連携して実施した。スマートフォン上で動作するアプリケーションの開発を容易にするため、MICSUS のサーバーシステムと連携する API をまとめたプラットフォームをスマートフォン側に構築し、そのプラットフォーム上で対話用アプリを動作させる構成とした。ぬいぐるみ型音声対話端末と同様の対話機能の実現に加え、電話がかかってくるなどで当該アプリが中断した場合でも、途中から対話を再開できる機能を実現した。これらの機能について実証実験を繰り返し行い、問題なく動作することを確認した。

## 2) マルチモーダル音声対話アプリと高齢者向けSNS システムの連携

高齢者向け SNS と音声対話アプリの連携のため、SNS コンテンツから対話時にシステムが発話し、オリジナルの SNS コンテンツへの導入となるセリフを抽出する、深層学習のモデルを構築した。NICT で開発している雑談対話システム WEKDA に統合し、MICSUS で SNS コンテンツを提示可能となる。

また、対象とする SNS コンテンツには写真が添付されていることが多く、そうした写真に対してユーザーが発する発話（例：「綺麗な花だね」、「昔、同じ場所に行ったことがあるよ」）に反応できる深層学習モデルも構築し、MICSUS で利用可能となった。

## 4 まとめと今後の展望

本研究開発を通じて、各要素技術のアウトプット目標を、また長期実証を通じて、トータルなシステムとしてのマルチモーダル音声対話システム（MICSUS）のアウトプット目標をそれぞれ達成した。さらに、生産性向上の側面では MICSUS を活用することで、面談 1 回に相当する情報収集に要するケアマネジャーの業務時間を、およそ 69% 低減できることを実証実験にて確認し、本研究開発の目的を十分に達成した。

実証実験の被験者（高齢者）からも、MICSUS に対する全体的な印象を聞くアンケートにて 5 段階評価の平均で 4.3 点と好印象の結果を得た（2022 年度に実施した高齢者 26 名による実証実験の評価）。具体的には、たまたま事前に楽しいことがあった被験者が実験に参加

した際、MICSUS が「今日は楽しそうですね？」と応答をした事例や、単にケアマネジャーが聞き取るべきことをヒアリングするだけでなく、Web 等の情報を用いた雑談に関して、MICSUS の雑談的応答の半分以上に対して、笑顔や「それは面白いですね」といった積極的な興味を示す反応が見られた等、好意的に受け止められたことなどがあり、貢献要因となっているものと考えられる。このように、被験者の状況を理解して寄り添った対話を行うことで、エンドユーザー自身が対話を楽しみ、飽きずに使い続けていただけそうな示唆を得ることができた。実際、実証実験に協力いただいたケアマネジャーからも、高齢者にも無理なく使っていただけるレベルまで改善してきており、十分にモニタリングの一助となり得る、とコメントをいただいております、ひっ迫する介護業界の現場からもその有用性を高くご評価いただいた。このようなことから、介護現場での生産性向上を成し遂げつつ、より高頻度（毎日等）かつタイムリーな健康状態のチェックにより介護の質の向上につながるとともに、疲れない AI が普段使いの会話相手としての役割を持つことで高齢者のコミュニケーションを促し、将来の健康状態の悪化リスクの抑制に貢献することが期待できる。

また、研究開発された MICSUS の技術を用いた新たなユースケースへの応用として、同 SIP 第 2 期のほかの研究課題と連携し、音声対話を活用したトイレでの状況確認、および地域包括ケアシステム「Team」への対話情報の連携についても取り組み、それぞれの目標を達成できた。

さらに、スマートフォン上で動作する SNS と連携した対話システムによって、高齢者のコミュニケーション不足の解消だけではなく、スマートフォンを介した多彩な協業の実現を後押しできると考えられる。

本研究成果については積極的に CEATEC などの展示会でも PR したことで、システムに関心を持つ自治体や介護業者などからも問い合わせが増え、実証実験に協力をいただくなど連携も進んだ。今後は、高齢者対話 AI サービスを検討している民間企業や社会課題・地域課題解決を目指す民間企業の支援（ライセンス提供など）、およびデジタルデバイドの解消を推進している地方自治体との共同実証など、ユースケースを広げて社会実装を進めていく予定である。

# 新しいEdTechプラットフォーム： 漫画リッシュによるテラーメイド英語学習

国立大学法人東京大学、エヌ・ティ・ティラーニングシステムズ株式会社

- 漫画を題材とした英語学習教材「漫画リッシュ」を考案し、英単語クイズおよび英作文演習を組み合わせ、テラーメイド英語学習プラットフォームを開発した。
- 漫画リッシュは、高校生に人気の漫画を英文穴埋め問題に解答しながら読み進める教材である。学習者は、漫画リッシュや英単語クイズの学習履歴に基づく英単語を使った、自由英作文の作成演習を行う。英作文は、英語で直接記入することも、DeepL<sup>\*1</sup>を使って日本語文の翻訳結果を解答欄に記入することもできる。
- 英作文の採点には、独自に開発したAI自動採点モジュールを導入している。学習者は、高得点を目指して繰り返し英作文演習を行うことで、英語表現力を向上させることが期待できる。

## 1 研究の目的

日本の国際競争力を維持・向上していく人材を確保していくには、グローバル化や情報化に対応した子供の育成が急務である。こうした背景にあって英語の表現力を効率的に伸ばす施策が課題となっている。本研究開発では、このような国際競争力強化につながるグローバルな人材を育成する観点から、英語を学習ターゲットとして選択した。

従来型の一斉型教育では、生徒の個別能力に応じた学習はできない。また、教科書の内容は日常生活で実際に使われる英語との差が大きく、教科書だけでは、生徒が日常会話の場面で使われるような生きた英語を多様な表現で使いこなせるようになるのは難しい。このため、従来の発想とは異なる、新しい英語学習法が必要である。

本開発研究では、主に高校生の英語学習をターゲットとし、新たな英語学習法を用いたテラーメイド学習の実現により、Society 5.0時代に求められる英語の学習効果を20%以上向上させることを目標とした。それら

の研究成果を社会全体に普及・展開していくことで、教育現場での新たな価値を創出する実用化例を数多く創出し、Society 5.0を具現化していく。

## 2 実施期間と方法

### (1) 実施期間

2018年度～2022年度

### (2) 実施方法

本開発研究では、高校生を対象とした英語学習法の研究に取り組むこととした。

研究開始にあたり、高校生をターゲットに実証実験を行うため、神奈川県横須賀市の三浦学苑高等学校を実証実験協力校として選定。学習ログを電子的にデータ収集するため、校内に無線LAN環境を構築し、実験対象者に学習アプリを搭載したタブレットを配布した（2年目以降は学校全体でタブレットを導入することとなったので初年度のみ）。

\*1 DeepL社が開発・提供している機械翻訳サービス。

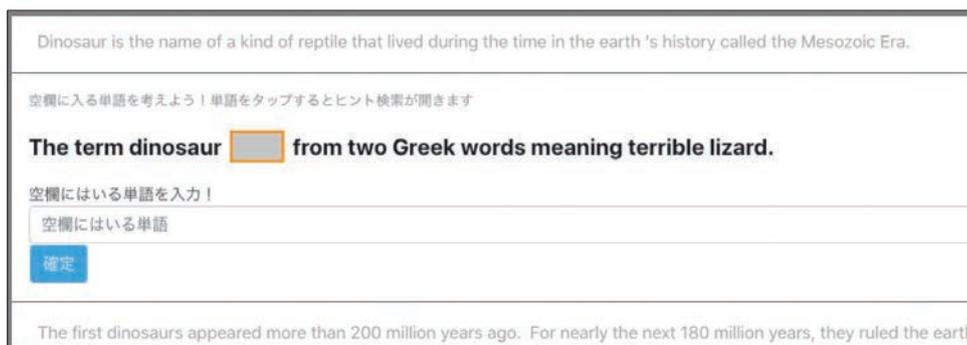


図1 ClozeTest 学習アプリの画面表示

三浦学苑高等学校では、実験用教材を用いた学習を授業の一部に取り入れてもらい、学習効果を検証した。また、放課後や夏休みなどの授業外期間に被験者を募集して実施する実験も行った。長期的な学力変化を把握するため、被験者となる学生は、NTT ドコモのE4S (English 4 Skills) の学力テストを年3回定期的に受験した。また、学習時の生体データの変化を研究することを目的とし、学習時のディープデータ\*2も収集した。具体的には、脳波計やNIRS\*3、視線計測を用いた学習時の生体データ収集や、日々の運動の活動量データ収集を行った。実証実験校は当初三浦学苑高等学校1校からスタートし、その後、明治大学附属明治高等学校やルネサンス高等学校グループなどを追加して規模を拡充した。また、より統制された実験を行うため、明治大学理工学部の学生を対象にした実験も行った。明治大学では、高校生大学生ともE4Sに加えTOEICを定期的に受験してもらい英語力の数値化を試みた。

英語学習法の検討にあたっては、Cloze Test という学習法に着目した。Cloze Test はところどころに空欄の空いた英文を穴埋めしながら読解する英語学習法で、学習者の言語能力を評価するテストとして有用なだけでなく、過去のさまざまな研究によって言語教育の教材としても有用性が期待されている\*4。

我々はCloze Test の効果を検証するため、タブレット端末上でCloze Test 問題を学習し、接触履歴や問題解答履歴といった学習ログを収集するCloze Test アプリを開発。2020年度より、三浦学苑高等学校の生徒を対象としたCloze Test アプリ学習実験を実施した。実験は1、

2年生全員を対象とし、Cloze Test アプリの操作法や学習方法の指導の授業を1時限実施した後、1年間自由に学習してもらう形式で実施した(図1)。

Cloze Test の長期的な学習実験を行う中で、二つの大きな課題が分かってきた。一つは学習者の離脱の問題、もう一つは表現力向上の問題である。一つめの離脱の問題は、Cloze Test アプリの出題の仕方に起因する。Cloze Test アプリは単純な英語の文章と空欄箇所を埋める問題で構成されており、継続して使ってもらうには飽きるという問題を克服する必要があった。二つめの表現力向上の問題は、Cloze Test の学習法に起因する。与えられた英文の一部を穴埋めするCloze Test の学習法は、与えられた英文を解釈し空欄箇所に入る適切な単語を考えるもので、自由な英語表現力が身に付くものではない。実際に英語が使用される場面は、ヒントとなる文などは存在しない。学習者に真に求められる英語能力は、自分が見たり聞いたり思ったりしたことが自由に表現できる英語表現力である。

離脱の問題に対しては、学習コンテンツとして高校生に人気のある漫画の英語版を活用することを考えた。漫画を活用することで、学習者は馴染みのコンテンツを楽しみながら学習できる。また、漫画コンテンツには日常会話で使用される「生きた英語」が多数使われており、教科書教材に比べて実践的な英語表現に多く触れられる利点もある。我々は、この漫画コンテンツを活用し、学習者が楽しみながら英語表現力を伸ばせる学習システムとして、漫画リッシュを活用したテーラーメイド英語学習プラットフォームを開発した。

\*2 その人の属性や趣味嗜好を深く長く蓄積したデータ。

\*3 近赤外線的光を利用して、脳や筋肉の血中のヘモグロビン濃度の変化を測定する装置。

\*4 Taylor, W. L.: "Cloze" readability scores as indices of individual differences in comprehension and aptitude., Journal of Applied Psychology, Vol. 41, No. 1, p. 19 (1957)



図2 漫画リッシュ実行画面と解説動画教材画面

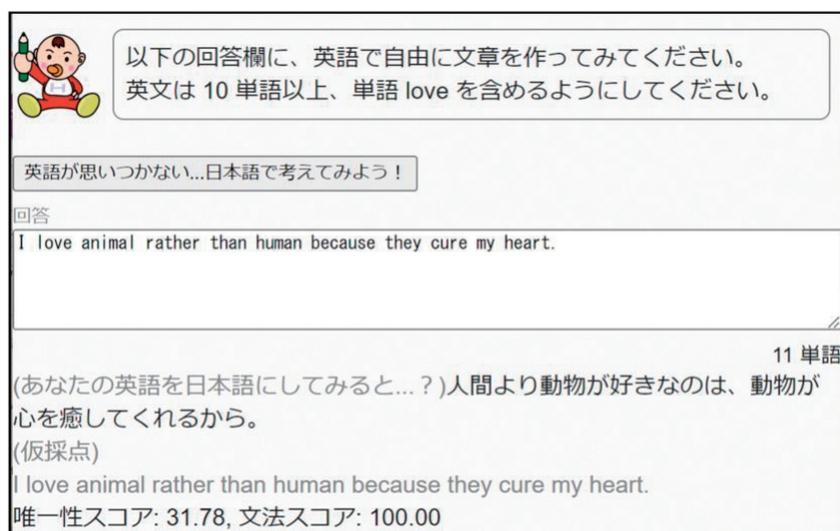


図3 英作文演習アプリ実行画面

### 3 研究の成果

漫画を題材とした英語学習教材「漫画リッシュ」を軸としたテラーメイド英語学習プラットフォームを考案し、パソコンやタブレット上で動作可能な学習アプリを開発した。漫画リッシュは、日本の著名漫画の英訳版を題材とした英語教材であり、吹き出しの一部が Cloze Test のように穴埋め問題となっている。辞書機能も搭載されており、学習者は分からない単語を辞書引きしながら漫画を読み進む。

漫画には、日常会話で使用される「生きた英語」が多数利用されている。そこで、漫画リッシュは予備校講師による生きた英語の解説動画教材を提供し、学習者に興味を持って英語を学んでもらうよう設計されている。学

習者は、漫画という慣れ親しんだ題材を読み、解説動画を視聴し、穴埋め問題を解きながら楽しく学習することで生きた英語を体得する（図2）。

漫画リッシュに加え、英単語クイズと英作文演習を加えた学習アプリ群を、テラーメイド英語学習プラットフォームとして教材化した。英単語クイズアプリは、NTT コミュニケーション科学基礎研究所（NTT CS 研）で作成された単語親密度データに基づき、学習者が学習すべき英単語を予測して出題する。英作文演習アプリは、課題として出された英単語を含む文章を学習者が自由英作文するアプリである。英作文は、英語で直接記入することも、DeepL を用いて日本語で作成した文章の翻訳結果を解答欄に入力することもできる（翻訳結果のコピーペーストは不可）。課題となる英単語には、漫画リッシュや英単語クイズで問題解答したり、辞書引きした単

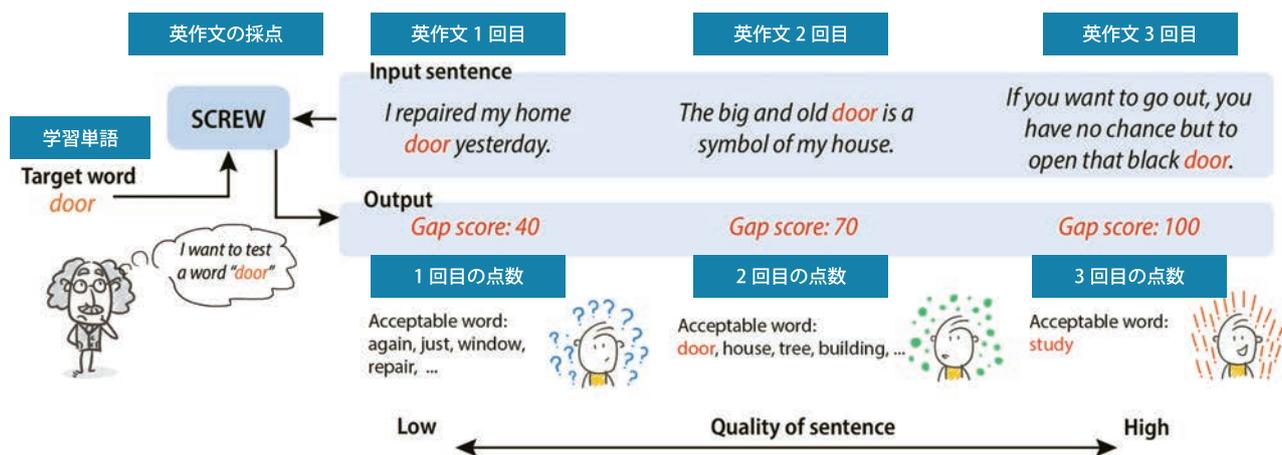


図4 SCREW の概念図

語を使用する (図3)。このように漫画リッシュで学習 (入力) された単語を、英作文で表現 (出力) する学習が、学習者の英語表現力向上につながることを期待している。

英作文演習アプリを学習者の表現力向上につなげるには、学習者がより多くの英作文を作成し、さまざまな英文に触れることが効果的であり、そのための有効な手段は、一つの設問に何度も英作文することである。そこで、AIによる英作文自動採点モジュール「SCREW」を開発し、英作文演習アプリに導入した。SCREWはAIを活用し、学習者が作成した英作文の表現力を測定するモジュールである (図4)。

表現力の測定は、BERTという自然言語処理モジュールのマスク言語モデルを用いた、空所単語予測を利用する。空所単語予測とは、文章中の一単語をマスクし、その箇所をBERTのマスク言語モデルと照合して、そこに入る候補の単語とその確信度を算出する機能である。SCREWは、空所単語予測を活用して出題単語の唯一性チェックを行う。具体的には、出題単語をマスクした際に得られる出題単語の確信度が高いほど出現単語の唯一性が高いと判定し、高得点の採点スコアを返す。SCREWはこのように測定することで、出題単語がほかの英単語に置き換え不可能かどうかをチェックして採点できる。ほかの単語で置き換え不可能な英作文は、出題単語を生かした表現力の高い英作文であることを示唆している。学習者は、SCREWの採点スコアを見ながら高

得点を目指して文章を練り直し、新たな英文を作成することでより多くの英文に触れることができ、結果として高い表現力が身に付く。

さらに、空所単語予測を活用した英文穴埋め問題自動作成モジュール「CLOZER」\*5を開発し、漫画リッシュアプリに導入した。CLOZERは、穴埋め箇所に入る候補単語の確信度がほかの候補より圧倒的に高い箇所を穴埋め問題とすることで、穴埋め問題の解答の曖昧性 (意図した単語以外の単語が正解として入ってしまう可能性) を抑制できる。CLOZERの活用により、漫画コンテンツ中にCloze Test問題を作成することが容易となる。

英語学習プラットフォームの学習状況把握や、学習者の学習モチベーション維持を図るため、それぞれの学習アプリを統合したダッシュボードを開発した (図5)。ダッシュボードには、各学習アプリの学習状況を可視化するだけでなく、個々の学習状況に応じた学習項目の提案に加え、学習ポイント制によるモチベーション維持の仕掛けが組み込まれている。学習ポイントは、学習者への「10ポイント獲得できたよ!」といった文言による動機付けを行ったり、自己ベストや他者と競争することでランキングによる動機付けが行える。

漫画リッシュをはじめとするテーラーメイド英語学習プラットフォームは、「GO-E-MON」という東京大学で開発した学習アプリケーション基盤システム上の学習アプリとして動作する (図6)。GO-E-MONは、分散PDS (Personal Data Store) を利用したオンライン実験プラ

\*5 松森 ほか. マスク言語モデルを利用した Open Cloze 問題の自動生成, 第 36 回人工知能学会全国大会, 3N4-GS-10-04 (2022)

\*6 S.Yazawa et al. GO-E-MON: A New Online Platform for Decentralized Cognitive Science, Big Data and Cognitive Computing, Vol.5, issue 4 (2021)

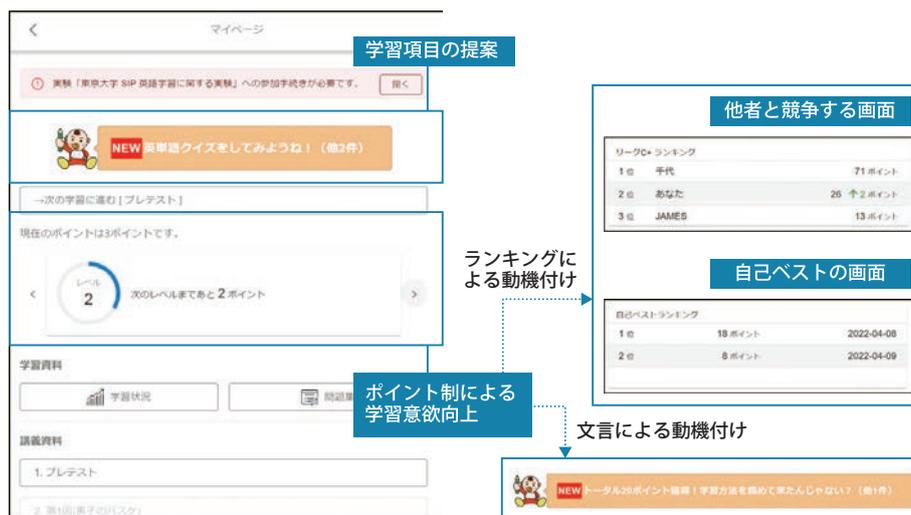


図5 ダッシュボードの画面イメージと機能

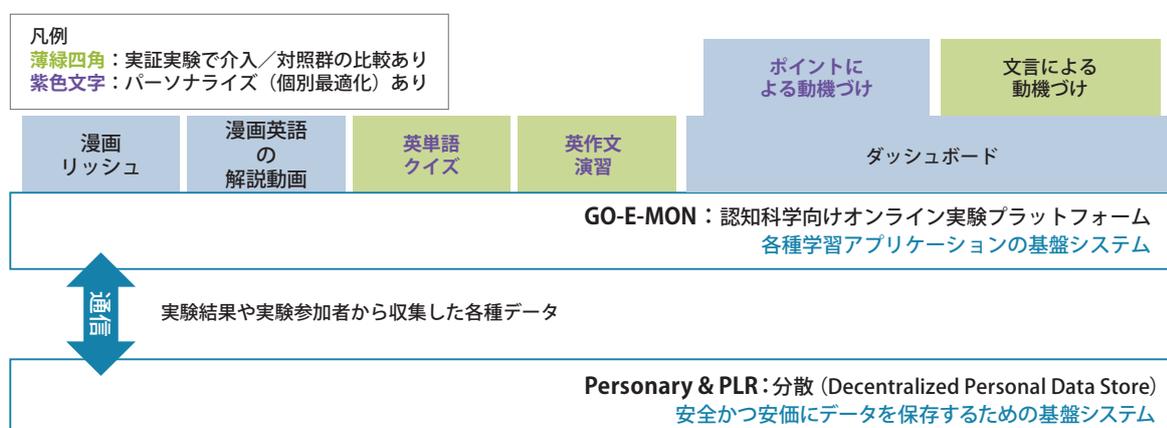


図6 英語学習システムの構成図

ットフォームである\*6。GO-E-MONを活用することで、漫画リッシュの実証実験環境を容易に構築できる。それぞれの英語学習アプリを学習した際に得られる、接触履歴や問題解答などの学習ログ、および学習者の個人情報は「Personary & PLR」という分散PDSに保存される。Personary & PLRでは、学習ログの所有権は学習者に帰属するとの考えのもと、学習ログの秘匿性やプライバシー保護を考慮してデータを保存する。

漫画リッシュを使った英語学習プラットフォームの学習効果を測定するため、プラットフォームを使った学習が英語力向上にどの程度結び付いているかを検証する実験を行った。実験は明治大学理工学部の学生19名を対象として行われた。被験者は、英作文演習の課題となる単語を漫画リッシュで辞書引きやクイズにより学習した接触単語を使用する群(随伴型10名)と、非接触単語を使用する群(非随伴型10名、1名ポストテスト未実施)の2つの群に分けられた。実証実験は、1ヶ月間のプラットフォームを使った学習記録と学習前後のTOEICの

スコアに基づいている。図7に漫画リッシュ学習前後における随伴型、非随伴型別のTOEICスコア分布の変化を表す箱ひげ図を示す。図中の点線で結ばれたプロットは同じ学習者の学習前(pre)、学習後(post)のTOEICスコアである。学習前に対する学習後のスコア増加率は、随伴型15.2%、非随伴型6.6%で随伴型の方がスコア上昇率が大きい結果が得られた。TOEIC等の一般的試験では、介入前の成績が高いと差分が表れにくい。これを考慮し、図8に学習前のTOEICスコアが平均的とされる600点未満の被験者13名に絞ったスコア変化を示す。この被験者群においては、スコアの上昇が顕著に見られ、特に随伴型において当初の学習効果の目標20%を超える21.8%の増加率が得られた。非随伴型の増加率は8.2%であり、この被験者群においては随伴型のスコア上昇率が明確になっている。

また、教員側の生産性向上に関する評価も実施した。三浦学苑の英語教員(8名)に唯一性を担保した英文穴埋め問題を実際に作成してもらった実験を行ったところ、

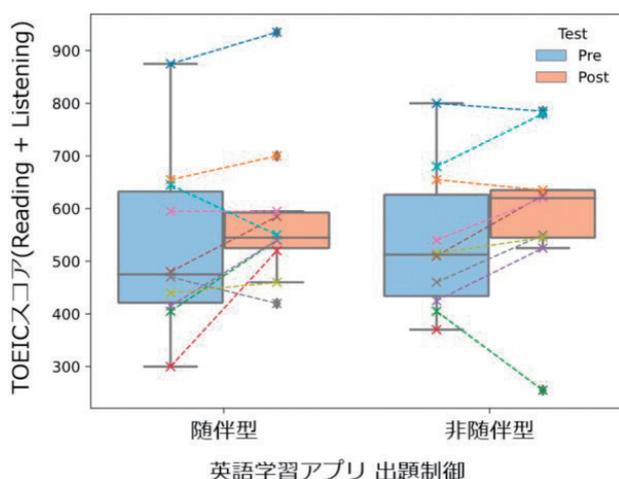


図7 英作文演習の出題制御別 TOEIC スコアの変化

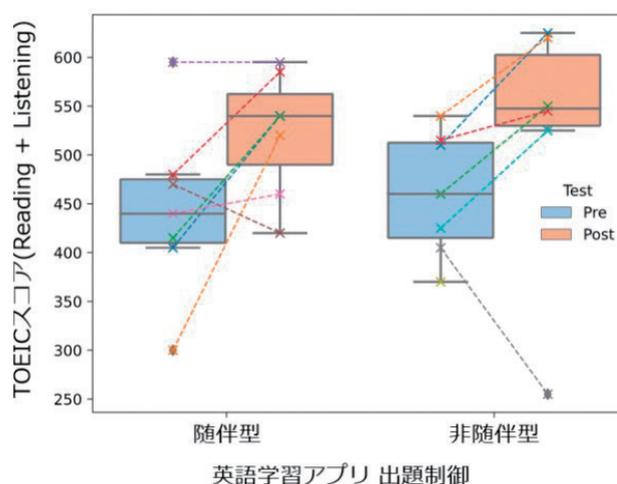


図8 英作文演習の出題制御別 TOEIC スコアの変化 (TOEIC スコア 600 点以下被験者。随伴型 7 名、非随伴型 6 名)

1 問作成に費やす時間は平均 240 秒であった。また、自由英作文の採点について教員に採点方法のヒアリングを行い、1 問の採点に 180 秒程度かかるとの結果を得た。これに対し、AI による英文穴埋め作成や英作文採点にかかる時間は最大 10 秒であり、コストは英文穴埋め問題作成 1/24、英作文採点 1/18 に改善していると考えられる。AI 導入による教員コスト削減の点に関しては、実行時間の短縮より問題作成や採点にかかるコストが 0 になることが重要である。例えば、一クラス 40 名の英作文採点を自動化することで  $3 \times 40 = 120$  分の教員コスト削減を実現することが可能となる。また、授業遂行面では、自由英作文の学習において SCREW が英作文を自動採点することで生徒が何度でも英作文にチャレンジできるメリットが得られる。同様の学習を教員が授業中に実施することは困難であり、AI 導入で新たな形式の授業が実施できることを示唆している。

高校生を対象とした実証実験では、被験者募集に応募のあった 1000 名程度の高校生に漫画リッシュプラットフォームを 1 か月程度自由に学習してもらい学習ログの分析を行った。現状の分析では、明治大学での実証実験と同様の結果が期待できる。ただし、学習意欲等をさらに高めるため、難易度の調整等を行う必要がある。

## 4 まとめと今後の展望

新しい英語学習の教材として、漫画を活用した学習教材漫画リッシュを考案し、東京大学で開発した学習アプリケーション基盤システム上の英語学習アプリとして実

装した。漫画リッシュは、英語版の漫画を題材とし、英文穴埋めクイズや講師による解説を交えて漫画を読み進める英語学習アプリであり、生徒は楽しく「生きた英語」を学べる。また、学力レベルに適応した英単語クイズや学習した英単語をもとにした英作文が演習でき、これらを繰り返すことで英語力の向上が期待できる。大学生を対象とした実験では、学習前 TOEIC スコア 600 点未満の被験者群において学習後 TOEIC スコアが 20% 以上向上することが確認できた。

また、漫画リッシュでは、AI による英文穴埋め問題自動生成モジュールと英作文問題自動採点モジュールを開発・導入し、英文穴埋めや英作文採点にかかる教員コストを削減することに成功した。これにより、漫画リッシュ導入によって教員コストを低減することができ、教員の生産性向上に寄与することが可能である。

今後は、さまざまな高校での実証実験による検証を続けるとともに、漫画リッシュのさらなる展開、ひいては事業化に向けた施策を行っていく予定である。コンテンツ増強に向けた出版社やコンテンツ事業者との連携、本教材の事業化パートナーとなる企業を選定、もしくはスタートアップ企業立ち上げの検討を進め、2023 年度末を目標に英語学習アプリの提供を目指す予定である。

また、本テーマでは、オンライン学習アプリの接触履歴や設問解答などの学習ログに加え、授業中の現場撮影や脳活動、活動量データ（運動データ）といったディープデータを多数収集している。これらの膨大なデータを活用していくことで、従来明らかになっていない学習効果の科学的な解明につなげることを目指していく。

# デジタル教材配信システムBookRollを用いた 教育データ収集・分析基盤システムと エビデンス共有システム

国立大学法人京都大学

- ラーニングアナリティクスとエビデンスに基づく教育を支援するために、学校教育現場のビッグデータ収集・分析基盤システムを構築し、エビデンスを蓄積・共有する仕組みを研究開発した。
- 中・高等学校の2教科（数学・英語）を対象とし、授業や家庭学習での学習者のスタディ・ログを活用する事例とその有効性を明らかにした。
- 知識マップに基づくAI推薦やデータ駆動の自己主導学習支援、デジタル読書による読解支援などの技術によって、エビデンスに基づくテラーメイド教育を検討した。

## 1 研究の目的

学校教育のあり方を Society 5.0 時代の学びの場へと進化させるため、ベテラン教師の経験や教育スキルをAI技術によっていつでも再現可能とし、エビデンスに基づく個人（学習者）の特性に合わせた、テラーメイド教育を実現する。具体的には、学校教育現場のビッグデータを取得・蓄積してAI技術と組み合わせることで、既存の学習方法・指導方法を解析・最適化し、学習者の個性や個々の習熟度に合わせた最適な学習コンテンツを提供するシステムを開発した。あわせて、教師と学習者、学習者とAIの双方向のインタラクションをリアルタイムに解析し、学習者の理解度や自主学習力、メタ認知力などを教師にフィードバックするシステムを開発した。

本研究開発では、教師・学習者間のインタラクションをスタディ・ログ化する研究開発として数学・英語を主な対象とし、授業や家庭学習での学習者（小・中・高等学校の生徒を想定）のスタディ・ログを収集・蓄積する際の収集データ項目と、その有効性を明らかにした。

また、学習支援による理解度変化を類型化し、有効な学習支援を実現する研究開発として、ペダゴジカル（教育学）情報プラットフォームに蓄積されたスタディ・ロ

グを分析し、認知科学的アプローチから類型化した。それによって、一人一人の学習者ごとに理解度や思考の特性に応じた学習支援（方法、内容、頻度やタイミング）が行える仕組みを研究開発した。

## 2 実施期間と方法

### (1) 実施期間

2018年度～2022年度

### (2) 実施方法

我々はラーニングアナリティクスとエビデンスに基づく教育を支援するため、「LEAF（Learning and Evidence Analytics Framework）」システムを開発した（図1）。この中心にあるのが、学習管理システム「LMS（Learning Management System）」で、教員や学生がパソコンやタブレットなどの情報端末を用いてシステムにアクセスし、学生の出席管理やレポート提出、成績管理などを行う。例えば、多くの大学では、MoodleやCanvas、SakaiなどのLMSを導入している。京都大学ではLMSと連携して教材を配信する「BookRoll\*1」というシステムを開発した。デジタル教科書や教師がつくった問

\*1 教員が教材を登録すれば、学生がブラウザで教材の閲覧ができ、学習ログが記録できるシステム。

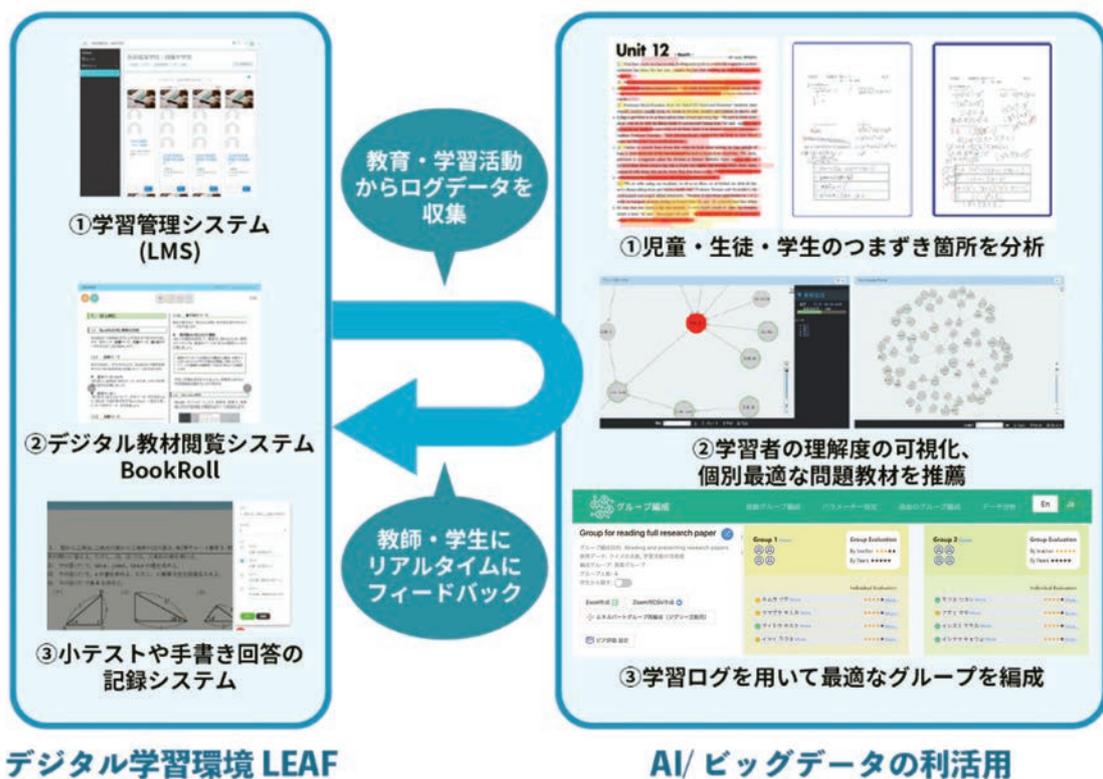


図1 LEAFシステムの全体像

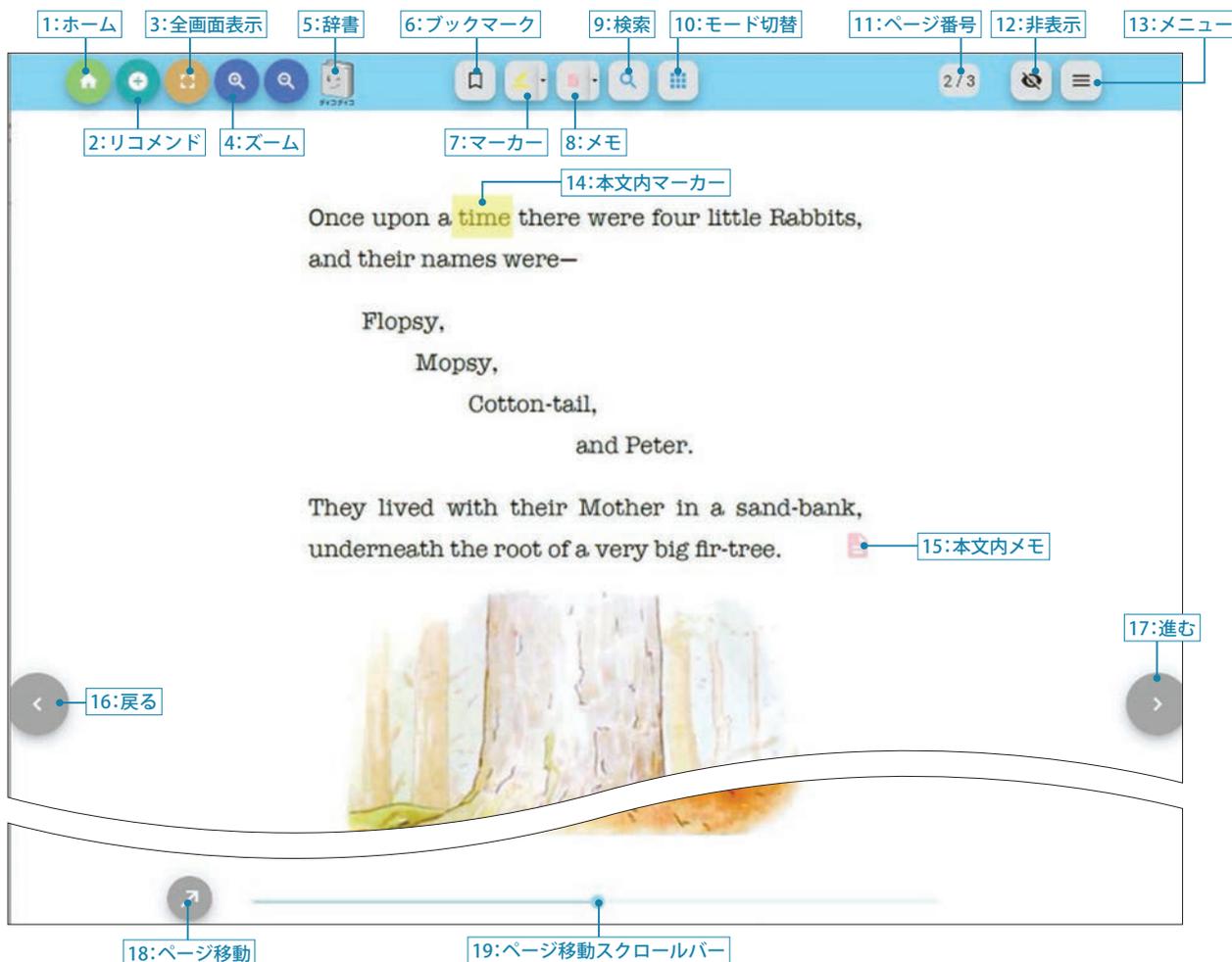


図2 教材配信システム (BookRoll)

題集および説明資料を登録し、学生はBookRoll上で教材を閲覧する。その際、閲覧履歴がxAPI（Experience API）\*2形式でラーニングレコーディングストア(LRS)\*3に蓄積される。BookRollに加え、LRSに蓄積されたデータを分析・可視化するツールとして「ログパレット\*4」というダッシュボードも開発しており、一つの学校に一つのLRSがあって、成績や日々の学習教育活動のデータが蓄積されることを想定している。ほかの学校にもLRSがあり、学生が小・中・高・大学と成長するに従い、分散したLRSを連結するためにブロックチェーン技術を応用する研究を行った。

BookRollは、Webブラウザで動作する教材配信シス

テムでありマーカー、メモ、クイズ、音声の再生などの機能を持つ（図2）。デジタル教科書や問題集、教師がつくったスライド教材や資料はPDFで登録できる。また、学習ログ収集機能を持ち、ページ移動、マーカー、メモ、手書き解答、クイズの解答などのログが収集される。

ログパレットは、LRSに蓄積されたスタディ・ログを分析可視化し、学習者と教師にフィードバックするダッシュボードである（図3）。例えば、基礎分析として、教材の閲覧達成率・時間、マーカーやメモの数、その内容、クイズの解答数・正解率などの指標が表示される。応用分析としては、教材の上に引かれたマーカーの一覧表示、手書き解答の解析、各科目の学習要素の構造と理解度の

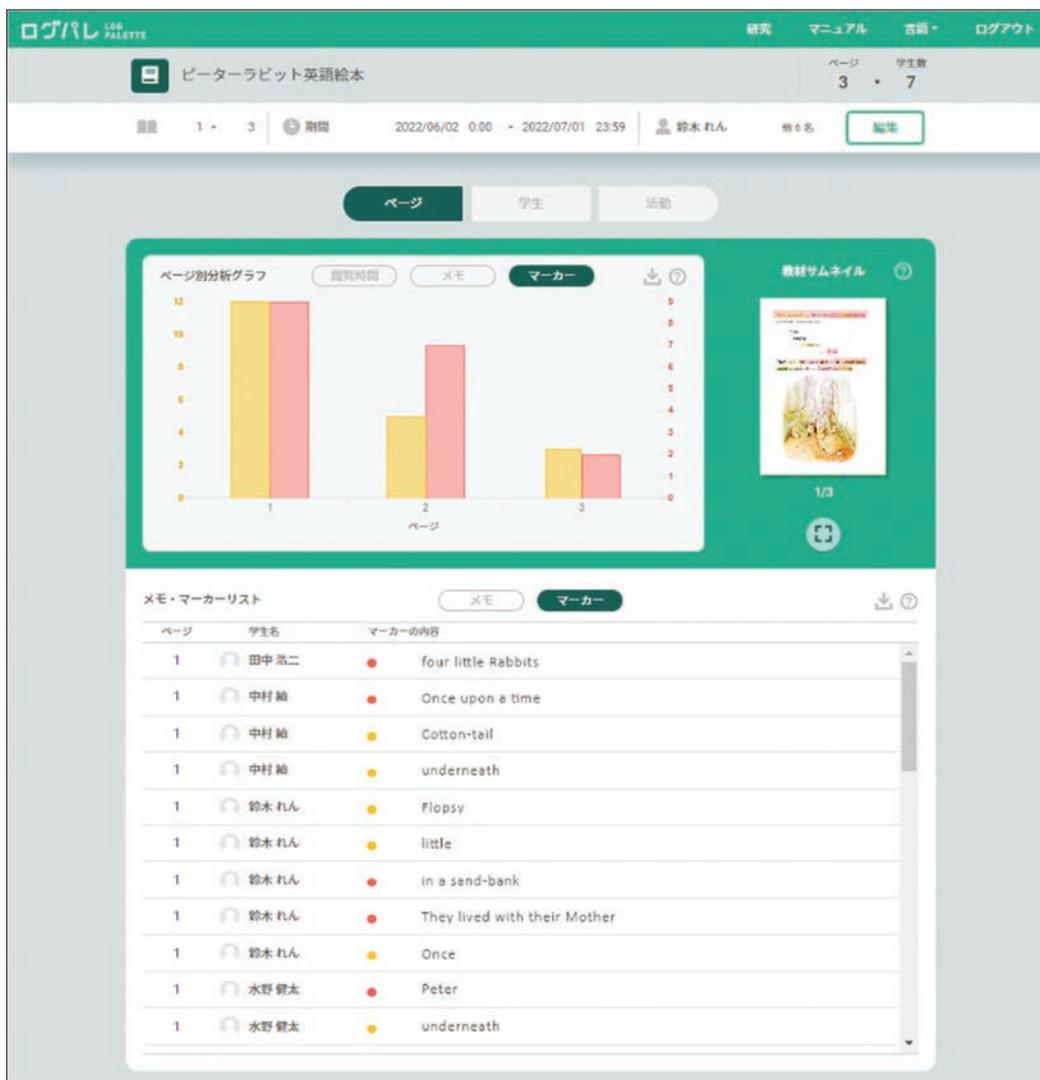


図3 スタディ・ログを分析可視化するダッシュボード(ログパレット)

- \*2 あらゆる学習履歴（経験）データを収集・記録するための、データ仕様のオープン標準規格。
- \*3 学習履歴データを格納するデータベース。
- \*4 蓄積された学習履歴データが分析可視化でき、学習者と教師にフィードバックできるシステム。

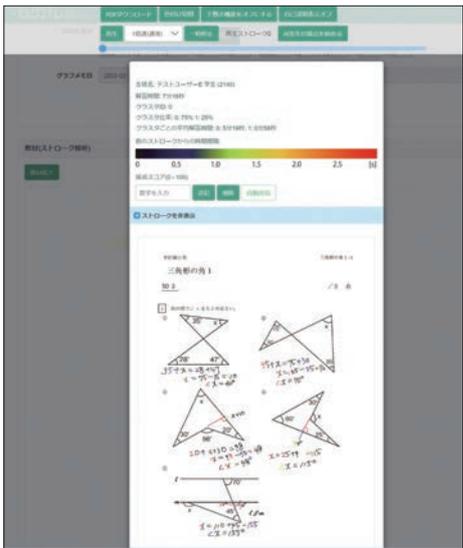


図4 つまずきの自動検出のための手書き解答解析



図5 自己主導的な多読学習支援 (GOAL)

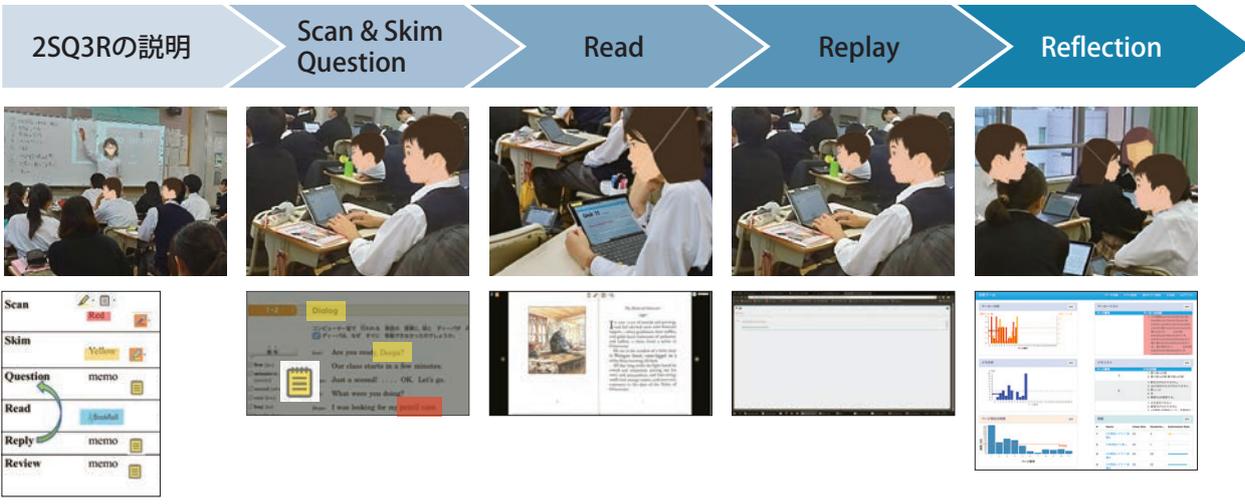


図6 アクティブ・リーディング戦略による読解支援

可視化（知識マップ）などが存在する。さらに発展モジュールとして、AI 問題推薦、AI 多読推薦、自己主導学習支援 (GOAL) などの機能がある。

**(3) 実証実験**

以下の実証実験を、京都市内の高等学校・中学校にて行った。

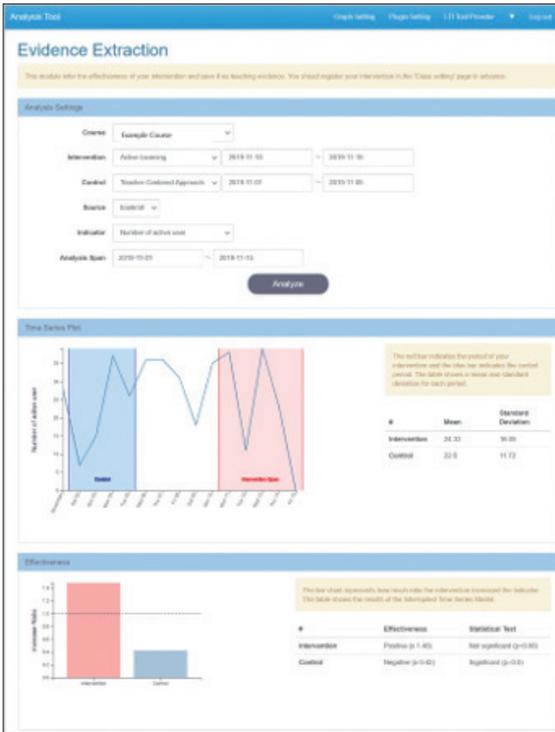
**①知識マップ・AI 問題推薦／手書き解答解析**

数学においては、中学1～3年と高校1年に知識マップを導入し、理解状況の推定とAI問題推薦機能を実装した。AI問題推薦のため、合計8,994問の数学演習問題をBookRollへ登録した。英語においては、多読用デジタル絵本約550冊を登録し、語彙・文法知識マップ作成のためのメタデータを抽出して中学・高校英語の語彙知識マップを構築した。そして中学生を対象にした

知識マップを用いた単語習得状況に基づき、多読用絵本の自動推薦を実施した。また、生徒のストロークの時間間隔に応じて、6段階のストローク色付け、解答過程のアニメーションを表示し、解答時間やストローク数、六つの色の割合などの特徴量に基づいてクラスタリングする機能を開発した（手書き解答解析、図4）。クラスタリング結果から典型的な解答を抽出して生徒にフィードバックする機能を導入し、中学（生徒360名）の数学で効果を検証した。

**②自己主導学習支援 (GOAL)**

生徒の自己主導学習能力を向上するため「GOAL (Goal Oriented Active Learner)」を開発し、状態分析や目標設定、モニタリング、振り返りを支援する機能を実装した（図5）。中学・高校の生徒計740名（高校1年生：



エビデンスの蓄積



エビデンスの登録

図7 エビデンス蓄積・共有のためのエビデンスポータル (REAL)

280名、高校2年生:100名、中学1～3年生:360名)が、自己主導的な英語多読を実施した。また、運動・健康活動を測定できるガーミンスマートウォッチを、中学1～3年生と高校2年生計460名に配布した。ガーミンスマートウォッチとGOALを通じて、自己主導的睡眠改善と運動チャレンジの検証実験を実施した。

③アクティブ・リーディング戦略による読解支援

高校1年生78名と高校2年生62名の英語授業でBookRollを使用し、アクティブ・リーディング戦略の手順として、ScanやSkim、Question、Read、Reply、Reflectionといった手法を導入して読解を進める2SQ3Rに沿って学習し、効果を検証した(図6)。

④エビデンスポータル (REAL)

エビデンスを蓄積・共有するための場所として、エビデンスポータル「REAL (Real-world Evidence Analysis Library)」を開発し、教師や研究者が自身の介入の内容、およびデータ分析方法について記入できるフォーマットを用意した。また、システムによる自動的な介入の効果推定を行う準備として、課題の内容や教授方法などの授業情報を収集するダッシュボードを整備し、授業モデル

の構築・検証を行った(図7)。

3 研究の成果

(1)実証実験①：知識マップ・AI問題推薦／手書き解答解析

1)知識マップ・AI問題推薦

中学・高校数学に知識マップとAI問題推薦を導入した結果、閲覧行動や知識マップ、コーススケジュールによる、クイズ推薦の個別最適化が可能になった。中学英語に知識マップと多読用絵本の自動推薦を導入した結果、多読活動の利便性と有用性の向上に加えて、生徒のモチベーションの維持や向上にも良い効果が得られた。今後は数学と英語学習を中心として、なぜ推薦されたかその理由説明を生成することでより納得して学習に取り組めるようにしていきたい。

2)手書き解答解析

手書き解答解析を3か月間実施し、教員2人にアンケート調査を行ったところ、システムの分かりやすさ、使いやすさ、今後の利用意欲に関して高い評価(5点満

点の5点)を示したことを確認した。また、教員2名へのインタビューでは、特に成績不振な生徒に有用であるとの意見を得た。今後は、ペンストロークから得られる学習プロセスデータを用いてつまづきポイントの検出やつまづきポイントに応じた推薦に応用することで、個別最適な学習支援を実現していきたい。

## (2) 実証実験②：自己主導学習支援 (GOAL)

### 1) 英語多読と自己主導学習支援

高校1年生47名を対象に3週間の自主多読を実施し、自己主導学習支援「あり・なし」の2グループで比較した結果、支援ありグループ(6,953単語数)は支援なしグループ(1,794単語数)に比べて多くの英語絵本を読んだ。中学1年生120名を対象に2か月間行った検証実験では、生徒が自分の多読目標を持っていれば、検証期間中に4時間以上の多読時間を確保でき、多読モチベーションを維持しながら自己主導学習能力を高めたことを確認した。さらに、中学1年生120名を対象とした1年間の実験を実施し、自己主導的な英語多読による訓練前後での多読スピードを比較したところ、全体で1分間あたりに読める単語数は99語から132語まで上がったとの結果が得られた。

### 2) 運動・健康活動と自己主導学習支援

睡眠改善の検証実験では、早寝早起き型の生徒の読書完了率が高くなり、学年が上がるとその関係性が強くなる結果が得られた。また、運動チャレンジの検証実験では、生徒の目標の大小ではなく、進捗のモニタリングと振り返りが歩数に大きな影響を与えることが明らかになった。今後は、睡眠改善・運動チャレンジのダッシュボードを開発して、学習者自身のライフスタイルを振り返って、自己主導能力を向上できることが期待される。

## (3) 実証実験③：アクティブ・リーディング戦略による読解支援

アクティブ・リーディング戦略による読解支援を検証した結果、生徒の語彙や読解力を向上させ、生徒の授業活動への参加促進に役立っていたことが明らかになった。さらに、アクティブ・リーディング戦略を活発に行った生徒は、不活発な生徒より英語の成績が高かったことを確認した。今後は読解学習を振り返ることの重要性を考慮し、アクティブ・リーディングに特化したダッシュボードを開発する。ダッシュボードを利用して学習者

が自身の読解成果を振り返り、授業の中で学習活動の一つとして共有することにより、学習者の読解力と読解意欲、メタ認知力等を高めることが期待できる。また、教師にとっては、学習者とクラス全体のログを確認することで、授業準備や次に取るべきアクションの判断などが効率的・効果的にできることが期待される。

## (4) 実証実験④：エビデンスポータル (REAL)

高校1年生の数学を対象に1か月間行った検証では、提案した授業モデルによって約65%の生徒が、普段よりも質の良い授業になったとの結果が得られた。今後は、システムによる自動的な介入の効果推定技術を導入し、フィードバックなどの介入の効果をエビデンスとして登録できる機能を開発する。この授業モデルとエビデンスポータルを用いて、蓄積されたエビデンスを用いたテーラーメイド教育を実現することが期待できる。

## 4

## まとめと今後の展望

本報告では、デジタル教材配信システム「BookRoll」を用いた教育データ収集、分析基盤システム「ログパレット」とエビデンス共有システム「REAL」について実証実験を行い、LEAFシステムを用いた学習者のスタディ・ログを利活用する事例成果を得た。各事例の成果と有効性により、ラーニングアナリティクスとエビデンスに基づく教育を実現する仕組みを構築した。

今後の展望として、実証実験①から④までのスタディ・ログ活用効果検証結果は、エビデンスポータルを通じてエビデンスレコードストア(ERS)に登録し、エビデンスを抽出・共有する仕組みを検証する。

また、一般社団法人エビデンス駆動型教育研究協議会(EDE法人)を通じて、政策立案者や研究者、実践者、民間企業等が連携して、当該システムを社会に展開する予定である。

# ペダゴジカル情報プラットフォームの実現と 社会実装に向けた研究開発

エヌ・ティ・ティ・コミュニケーションズ株式会社、日本電信電話株式会社

- ベテラン教師の経験や教育スキルを AI 技術によっていつでも再現可能とし、エビデンスに基づいて、個人（学習者）の特性に合わせたテーラーメイド教育の実現を目指す。
- 教員や塾、ベンチャー企業、教育行政担当者などにも共有可能なペダゴジカル（教育学）情報プラットフォームについて研究開発を行った。

## 1 研究の目的

学校における教育のあり方を Society 5.0 時代の学びの場へと進化させるため、ベテラン教師の経験や教育スキルを AI 技術によっていつでも再現可能とし、エビデンスに基づいて、個人（学習者）の特性に合わせたテーラーメイド教育を実現する。具体的には、学校教育現場のビッグデータを取得・蓄積して AI 技術と組み合わせることで、既存の学習方法、指導方法を解析・最適化し、学習者の個性や個々の習熟度に合わせ、継続的な学習に追従して最適な学習コンテンツを提供するシステムを開発した。

また、スタディ・ログの収集や蓄積、分析を行うために、適切な加工や同意などを許諾した上で、教員のみならず、塾やベンチャー企業、教育行政担当者などにも共有可能なペダゴジカル情報プラットフォームについて研究開発を行った。

## 2 実施期間と方法

### (1) 実施期間

2018 年度～ 2022 年度

### (2) 実施方法

1) 個別最適化された学習支援を実現する「ペダゴジカル AI エンジン」の研究開発

2020 年度は、英語学習のためのインタラクティブコンテンツ提示環境として、①英語多読に適した教材提示・英文読み上げ・確認問題生成を行う英語学習支援システム「Enavis」の機能を拡張し、Listening 学習用音声データ生成・語彙／文法項目などに関連する穴埋め問題自動生成を行うとともに、②英語発話リズム訓練アプリの機能を拡張し、Speaking の質が向上する（発話リズムスコアが平均 80 以上となる）ような訓練方法の構築と実証を行った。

2021 年度は、

- ・英語学習者の語彙力レベルを推定する語彙数推定技術
- ・英語絵本や英字新聞記事などの著作権処理済みの英文コンテンツを提示する英語多読学習支援システム「Enavis」
- ・英文コンテンツからの文法問題自動生成技術
- ・本人の発話音声でよりネイティブの発話リズムに近い音声に変換する英語発話リズム訓練アプリ

の四つを連携させた、インタラクティブ英語学習支援統合環境のプロトタイプを構築した。

2022 年度は、インタラクティブ英語学習支援統合環境を活用して英語発話リズム訓練アプリのスコアと、Enavis・語彙数推定・問題自動生成の結果の個人内での相関分析を行い、アプリによる Speaking の質の向上が

Listening、Reading、Writingの質の向上あるいは能力とどのように関連しているのか検証した。

具体的には、インタラクティブ英語学習支援統合環境で取得したデータを対象とし、理解度変化・推定技術の実証、および学習効果推定・予測技術、学習プラン生成・レコメンド技術を構築してそれぞれ効果を実証した。学習効果予測およびレコメンド結果については、上記インタラクティブ英語学習支援統合環境における問題自動生成や、次の学習プランを提示して学習者の学習行動を継続的に蓄積。その効果を、統合環境で取得される語彙数変化・小テストの成績と照らし合わせて評価し、効果を検証した。

また、実証実験においては、商用学習サイトE4S(English 4 skills)のコンテンツを教材として、ペダゴジカルAIエンジンによる問題推薦をあわせて実施した。

## 2) 相互運用性を確保したペダゴジカル情報プラットフォームの研究開発・実用化検討

2020年度は、プロトタイプLRS\*1の有用性を実証校で検証し、LRSをNTTコミュニケーションズのサービスとして提供できるようにした。

2021年度は、コンテンツ提供事業者と交渉し、NTTコミュニケーション科学基礎研究所(CS研)などが行うペダゴジカルAIエンジンの研究開発に必要となるスタディ・ログを収集・提供できるようにした(CS研などと連携し、スタディ・ログ収集に必要なデータセットを明確化)。

さらに、ペダゴジカルAIエンジンの研究開発状況・実用化判断を踏まえ、自走するプラットフォームとしてのビジネスモデル案を策定した。

2022年度は、社会実装に向けて、AIエンジン自体の効果検証、市場性、接続に向けた技術検証、およびシステム連携方式を研究開発した。

## 3) 相互運用性が担保されるスタディ・ログの標準規格の策定

2020年度は、学習要素リストの標準策定などを目標とし、xAPI\*2プロファイル仕様案と算数・数学に関する学習要素リストをスタディ・ログ標準規格として策定

した。

2021～2022年度は、xAPIプロファイル仕様案の更新などを目標とし、2020年度に策定したスタディ・ログ標準規格(xAPIプロファイル仕様案と算数・数学に関する学習要素リスト)を、今後のLRSとのデータ連携の際に標準仕様として活用してもらえるようコンテンツ事業者に対して交渉を継続。xAPIプロファイル仕様案については、随時更新を実施中。

## 3 研究の成果

### (1) 個別最適化された学習支援を実現する「ペダゴジカルAIエンジン」の研究開発

2020年度は、多様なスタディ・ログデータ(特にテスト問題の正誤、欠損があってもかまわない)を対象として、潜在的な学力特徴を有する学習者群を教師なし機械学習の手法で効率的に抽出し、理解度の向上要因を推定する手法を構築した。具体的には、これまでに構築したMonotonic VAE法\*3を欠損値にも対応するよう拡張し、学習者間で異なるテスト問題に関するスタディ・ログを用いて、広範な学力調査データの分析を実現した。また、特徴的な学力を有する学習者群の抽出結果に対して、専門家の意見を取り入れながら有効な教材をレコメンドした。さらに、学習者の理解度変化データを蓄積することで、理解度変化学習者群および個々人の学力向上に有効な教材の抽出や、効果推定手法およびレコメンド手法のプロトタイプを構築し、スタディ・ログの活用と理解度変化に基づくテラーメイド学習環境構築に有効なエビデンスを抽出した(図1)。

2021年度は、50名程度の大学生を対象に1万語以上の英単語親密度データベースを構築し、少数の語彙のチェックによって学習者の語彙数を推定する技術および、学習者の学習履歴やテスト結果から未習得語の候補を推定する技術を構築した。インタラクティブ英語学習支援統合環境の機能検証においては、学習者数十名を対象に実証を行い、学習効果エビデンスを収集した。

また、2020年度に開発した理解度推定技術、学習効果推定技術について、標準学力調査などで取得される大

\*1 学習履歴データを格納するデータベース。

\*2 あらゆる学習履歴(経験)データを収集・記録するためのデータ仕様のオープン標準規格。

\*3 深層学習のモデルの一つ。

規模なデータ（試行用標準データ）に基づき、スタディ・ログの蓄積に対応した時系列データに拡張し、逐次学習の効果を捉える理解度推定技術、学習効果推定技術を構築した。さらに、試行用標準データに基づく問題を推薦した学習者群の、理解度向上の効果が見込まれる教材推薦技術を構築し、インタラクティブ英語学習支援統合環境をはじめ、各実証校で取得し得る多様なスタディ・ログに対応させ、個々の学習者に適したテーラーメイド学習のエビデンスを蓄積した。また、2022年度の実証に向け、インタラクティブ英語学習支援統合環境と連携する上での全体アーキテクチャ設計・構築を実施した（図2）。

2022年度は、高等学校の英語学習支援において、長期的にインタラクティブ英語学習支援統合環境を活用することにより、継続的にスタディ・ログを蓄積するとともに、動的な問題推薦を行い学習効果の検証を実施した。実験に際しては、プロトタイプシステムを改修し、インターネット経由で多人数が同時に利用できるシステムを構築した。また、先行実験への実験参加者（生徒および教員）からのフィードバックを受けて教材の改良を継続的に実施した。

先行実験で得られた継続的な学習データを用いて、ペダゴジカル AI エンジンの各要素技術である、①理解度変化・推定技術の実証、②学習効果推定・予測技術、③学習プラン生成・レコメンド技術を改良し、長期的なスタディ・ログから得られる個々の学習者の理解度変化に基づき、学習効果予測および問題レコメンドによるテーラーメイド学習の効果を検証する実証実験を実施した。

実証実験で収集したスタディ・ログにペダゴジカル AI エンジンの各要素技術を適用することにより、個々

の学習者の理解度の変化を潜在学力特徴空間上で時系列的に把握し、個々の学習者のその時の理解度状況に応じて、未学習のそれぞれの問題に対する予測正答率の高精度な推定を実現した。これにより、個々の学習者の理解度状況に適した難易度の問題群を自動的に抽出し、学習者の個々の潜在学力特徴を向上するのに有効な問題群を序列化することが可能となった。

実証実験では予測正答率75%の問題群を推薦し、ランダムに問題を解いた際の正答率が50%程度の群と85%程度の群のどちらに対しても、実際の正答率が目標値である75%程度に収束し、長期的な学習による理解度状況の変化に追従して、学習効果が期待できる問題群を動的に推薦できる機能を実証した。以上のペダゴジカル AI エンジンの各要素技術の機能実証により、学習の効率を向上するとともに、学習者のモチベーション維持に効果のある問題レコメンドが可能な学習環境を実現した。長期的な学習の効果については、実証校の協力の下で外部試験の受験等による検証を継続している。

## (2) 相互運用性を確保したペダゴジカル情報プラットフォームの研究開発・実用化検討

2020年度は、蓄積された学習ログの要素を、可視化用ダッシュボードによって、コンテンツメタデータやユーザー属性情報などの外部情報と組み合わせることにより、対象となる期間やコンテンツ、学年、クラスの組み合わせで選択した母集団に対して、「勉強時間」、「ユニーク利用者数」、「行動種別割合」、「利用時間帯」を可視化できることを確認した。

2021年度は、スタディ・ログの収集にあたって、LRSとして構造化されたデータに加え、非構造化データ

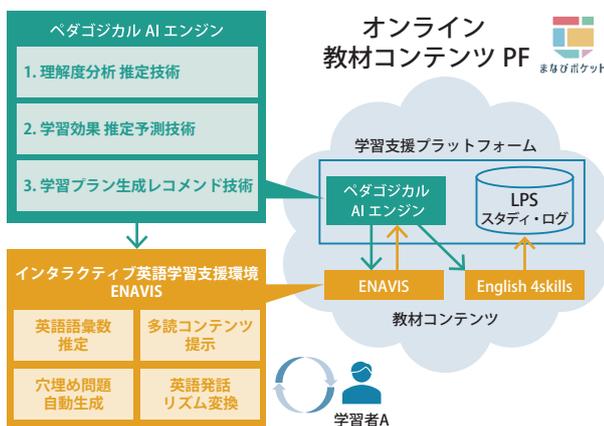


図1 テーラーメイド学習環境



図2 英語学習支援統合環境「ENAVIS」

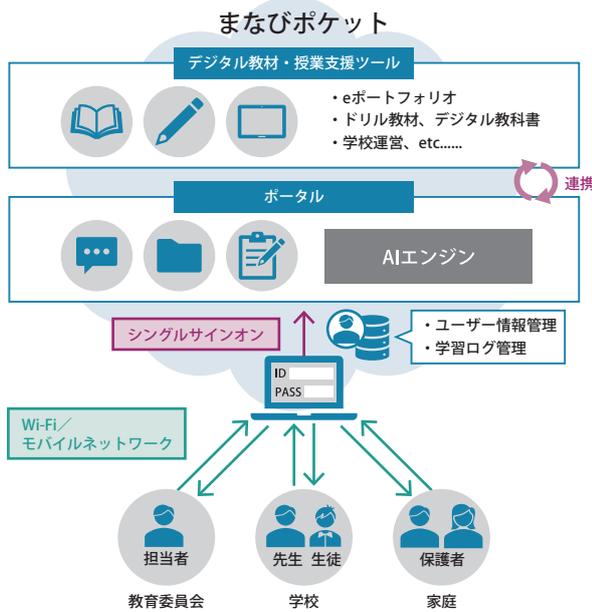


図3 ペダゴジカル情報プラットフォーム

を含めてデータを取得した。また、データ蓄積基盤であるLRSを、NTTコミュニケーションズの「まなびポケット」にて実装し、ドリル教材などを提供するコンテンツ事業者が横断的にデータを活用できる仕組みを構築した。

商用化に向けての性能要件対策については、LRS単体と外部情報連携のデザインパターンにそれぞれCQRS\*4と並列分散処理が適切であることを導出しこれを採用した。LRSにCQRSパターンは、まなびポケットの認証に基づいた検索クエリ制御が行えるため、教員や教育委員会といったロール制御や所属自治体によるデータ共有の制御が可能になった。

自走するプラットフォームとしてのビジネスモデル案は、NTTコミュニケーションズの全国の営業組織が提案・販売を担い、教育委員会や学校法人などへの直接販売、および販売代理店や地場ベンダーなどを経由した間接販売を想定している。

拡販に向けたプロモーション施策としては、

- まなびポケット公式 Web サイト、Facebook、Twitter での情報発信
- 教育委員会・教職員向け活用研修（個別／ウェビナー）
- 販売店向けサービス説明会
- セミナー開催
- 各種イベントへの出展

\*4 コマンドクエリ責務分離を表し、データストアの読み取りと更新の操作を分離するパターン。

カテゴリ	課題	対応策
AI エンジン 自体の効果 検証	<ul style="list-style-type: none"> <li>• 学習履歴の時系列を対象とする分析</li> <li>• 多種類の教材/学習履歴を対象とする横断的な分析</li> </ul>	<ul style="list-style-type: none"> <li>• 相互運用性を確保したペダゴジカル情報プラットフォーム上での実証実験の実施</li> <li>• 上記プラットフォーム上での一般利用者からの長期的な実証実験データの取得</li> </ul>
社会実装 (商用化に 向けて)	<ul style="list-style-type: none"> <li>• 市場性 サービスとして需要性、顧客ニーズの確認</li> </ul>	<ul style="list-style-type: none"> <li>• これまで弊社がまなびポケットで築いたリレーションを活用し、教育委員会/学校法人等、販売代理店への需要/ニーズ調査を実施</li> </ul>
	<ul style="list-style-type: none"> <li>• 技術面 AI エンジン搭載(商用化)に向けた要件整理、必要なリソース/インフラ設計/システム連携方式の検討</li> </ul>	<ul style="list-style-type: none"> <li>• CS 研究成果である AI エンジンを実装し検証を実施</li> </ul>
	<ul style="list-style-type: none"> <li>• 法制度面 個人情報等を含むデータを第三者にオープンにするための法制度への対応</li> </ul>	<ul style="list-style-type: none"> <li>• 事業者毎の対応ではなく、本実証で得られた知見や各種ガイドラインに沿った対応</li> </ul>

表1 社会実装に向けた課題

などを想定している。

### (3) 相互運用性が担保されるスタディ・ログの標準規格の策定

2020年度は、コンテンツ事業者に対して、今後のLRSとのデータ連携に向けた際の標準仕様としてxAPIプロファイルの活用を提案した。また、学習要素リストについては、複数(2社)のコンテンツ事業者と協力し、コンテンツ事業者の自前教材と突合いを行い有用性について検証した。

2021年度は、xAPIプロファイル仕様の随時更新を実施した。

## 4 まとめと今後の展望

最終形態は、NTTコミュニケーションズが提供する「まなびポケット」上にCS研が開発したペダゴジカルAIエンジンを実装し、同社の販路を活用して全国の教育委員会をターゲットに商用展開を行うことを目指している(図3)。

なお、社会実装に向けては、表1に示す課題への対応が必要と考える。

# 記号的AIに基づく思考経験のデザインと 統計的AIに基づく思考パターンの検出による テーラーメイド学習支援

国立大学法人広島大学

- 本研究開発では、記号的 AI と統計的 AI の双方を活用したテーラーメイドな学習支援の実現を目指した。
- 研究は、「①記号的 AI に基づく思考経験のデザイン」、「②実践を通して得られたデータに基づく統計的 AI を用いた思考パターンの検出」、「③思考パターンに基づくテーラーメイド学習支援の開発」、「④ポータブルでセキュアな分散型データ管理」という四つのサブテーマに分けて行った。
- 記号的 AI に基づく思考経験のデザインの有用性を、教育現場における実践を通して確認できた。また、実践を通して得られたデータに統計的 AI を適用することで、思考パターン検出の可能性が示せた。

## 1 研究の目的

本研究開発における「記号的 AI」とは、人の持つ知識を記号的に記述し、その記述をもとに推論する枠組みである。この枠組みを学習者の思考の枠組みとして提供し、学習者にその枠組みに沿った思考を経験させる場をコンピューター上で実行することが「思考経験のデザイン」である。デザインされた思考の場は、適度な自由度を保ちつつ限定された思考が行われることから、思考に関する、より直接的で構造化しやすいデータが取得できる。取得されたデータを用いて思考パターンを検出する際、統計的数理モデルをベースとする人工知能技術を用いるが、これを「統計的 AI に基づく思考パターンの検出」としている。

従来の学習支援が、解答の正誤や誤答の種類に基づいていたのに対して、思考パターンが検出できれば、「どういった思考を行ったのか」に基づき学習を支援できる。これは、より個々の学習者に合わせた支援となり、本研究開発では「テーラーメイド学習支援」と定義する。

## 2 実施期間と方法

### (1) 実施期間

2018～2020 年度

### (2) 実施方法

本研究開発では、「①記号的 AI に基づく思考経験のデザイン」、「②実践を通して得られたデータに基づく統計的 AI を用いた思考パターンの検出」、「③思考パターンに基づくテーラーメイド学習支援の開発」、「④ポータブルでセキュアな分散型データ管理」という四つのサブテーマを設定した。図1に示す研究体制スキームを構築し、②のテーマについては広島工業大学、③のテーマについては山口大学に再委託して共同で研究開発を行った。

本研究開発のベースとなる、「記号的 AI に基づく思考経験のデザイン」の枠組みを図2に模式的に示す。情報処理アプローチに基づく認知科学においては、人の思考はある側面から知識と推論によってモデル化できている。記号的 AI は、知識を記号化し、その記号処理として推論をコンピューターで実現することを目指したものである。

これに対して、認知的に妥当性を保って知識や推論を記号化できている場合、その記号を学習者に提供するこ

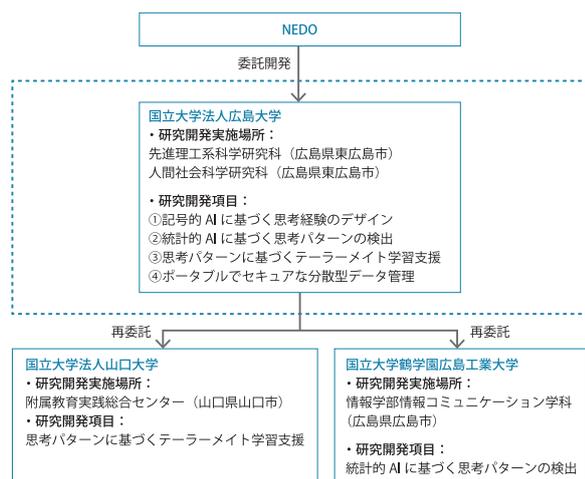


図1 研究体制スキーム

とが本研究開発となる。従来の記号的AI・エキスパートシステムに関する研究開発においても、コンピューターが行った推論や知識を人が理解できるように説明することは非常に重要な研究課題であり、統計的AIに対する優位性の一つであると考えられてきた。本研究開発では記号化された知識や推論を、学習のための外的資源・思考を外在化する道具として、より積極的に利用する。

本研究開発では、学習環境で提供された外的資源の操作として思考が外在化され、その操作データが思考のデータとなり、統計的AIの処理対象となる。統計的AIは近年非常に大きな成果を上げているが、対象データの前処理および出力結果の解釈が簡単ではなかった。この二つが定式化しやすい画像や音声、あるいは翻訳としての言語においては多様な研究が進められているが、これらの定式化が難しい対象に関しては、必ずしも成果が上がっているとはいえない。

本研究開発は、一般的にはデータの取得自体が難しく、また、関連データを集めてきたとしてもその前処理や解釈が難しい思考に関するデータ処理やパターン抽出を、思考の外在化、すなわち外的資源の操作データおよびその解釈として行う。思考パターンが抽出できれば、個々の学習者の持つパターンに基づいた学習支援が可能となる。従来の学習支援に関する研究が、学習者の解答の正誤や誤答の種類あるいは学習者のこれまでの正誤履歴といったものをベースとしていたのに対し、学習者の思考パターンに基づくことで、学習者の思考に適した個性

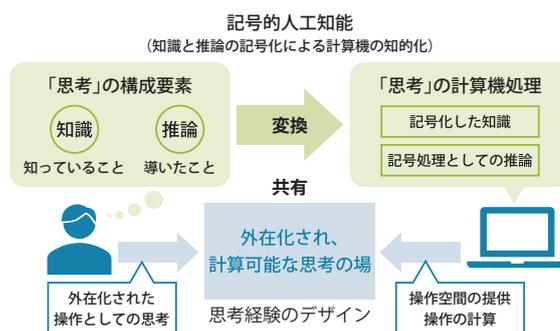


図2 思考経験のデザインの枠組み

の高い支援が期待できる。また、これらを大規模に実現する上で、多くの学習者から安全な形でデータを学習する仕組みを用意する必要があり、そのためのポータブルでセキュアな分散型データ管理の設計・開発も行った。

### 3 本研究の成果

#### (1) 記号的AIに基づく思考経験のデザイン

本研究開発での記号的AIに基づく思考経験のデザインとしては、これまで行われてきた先行研究開発を踏まえ、「単位算数文章題を対象とした組立てソフトウェア」、「複合算数・数学文章題を対象とした量間関係構造組立てソフトウェア」、「英文読解を対象とした概念マップ組立てソフトウェア」をそれぞれ実践的に整備し、教育現場で利用しながらデータを収集した。

##### 1) 単位算数文章題を対象とした組立てソフトウェア：モンサクン

図3に単位算数文章題を対象とした、組立てソフトウェア：モンサクン\*1のインターフェースを示す。このソフトウェアでは、学習者に対して外的資源としてカードの形で量概念\*2を与え、量概念の組合せとして問題を組み立てさせる。図3の場合、右に並んだ五つの項目は言語で表される一つの量概念を表している。個々の量概念は、量が属する「主体」、量としての「ラベル」

\*1 広島大学で開発された、「問題を解くのではなく、問題をつくる」ことで学習するソフトウェア。問題をつくる過程において問題の持つ「構造」を意識することで「構造を理解する力」を養う。

\*2 文章題に現れる量を表わす概念であり、数は量概念の値として意味を持つ。

および「値」の三つで構成され、実体（インスタンス）としての量概念となっている。この量概念を組み合わせると、概念間に演算関係が生じることがある。指定された演算関係が生じるように量概念を組み合わせることを目的に、文章題を組み立てさせる。具体的には、適切なカードを選んで左にある三つの空欄を埋めることで、カードの組合せをつくる。なお、単位文章題とは、演算一つの演算関係を生じる文章題のことである。

本研究開発では、2020年度にある小学校と協議し、全学年でのこのような文章題を授業で実施した。実施内容および収集したデータの概要を表1に示す。達成課題数、解答数/課題数、ステップ/課題数は平均である。ステップ数は、カード操作回数を表す。授業は617名の児童が参加し、総計1,040時限・人のシステム利用を行った。正しくつくられた問題数が31,583問であり、作問の総数が81,205問であることから、約50,000問の不適切な作問事例が得られ、280,664ステップのカード操作データが得られた。

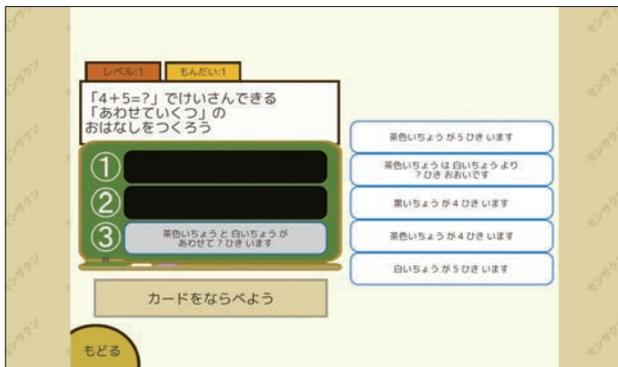


図3 単位算数文章題を対象とした組立てソフトウェア：モンサクン

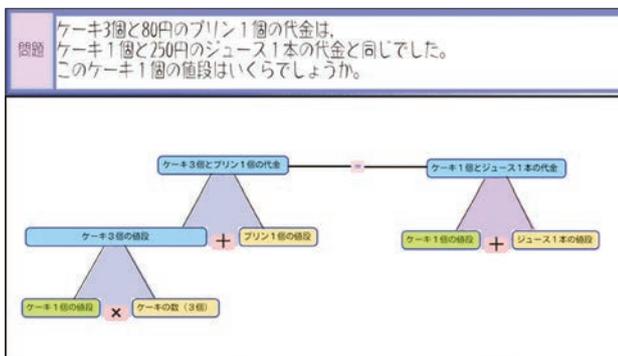


図4 組み立てられた三角ブロックの例

## 2) 複合算数・数学文章題を対象とした量間関係構造組立てソフトウェア：三角ブロック

図4に複合算数・数学文章題を対象とした、量間関係構造組立てソフトウェアの三角ブロック図を示す。学習者は、カードの形で提示された量概念の演算関係を構造的に組み立てた。組み立てられた演算関係を、既知の値を用いて順次計算していく場合は算数的な求値方法となり、この数量関係を方程式化し代数的に解く場合には数学的な求値方法となる。

ソフトウェアについては、一つの小学校において228時限・人（5、6年生228名、724課題、6,146解答、44,481ステップ）、二つの中学校において計392時限・人、一つの高校において計600時限・人の実践利用を行い、運用可能性とデータ収集可能性を確認している。

## 3) 英文読解を対象とした概念マップ組立てソフトウェア：キットビルド概念マップ

図5に英文読解を対象とした、キットビルド概念マップ\*3の組立て例を示す。このソフトウェアでは、読

教材	学年	時限	人数	達成課題数	解答数/課題	ステップ/課題
和差文章題 作問	小1	2	104	22.1	2.81	10.5
	小2	2	105	31.2	2.67	9.70
	小3	2	89	35.5	2.76	9.99
	小4	2	93	38.7	2.29	8.60
	小5	2	122	41.7	2.28	8.90
	小6	2	104	43.2	2.19	8.33
乗除文章題 作問	小3	1	83	51.2	2.63	6.88
	小4	1	87	62.3	3.03	9.28

表1 実践および収集データ

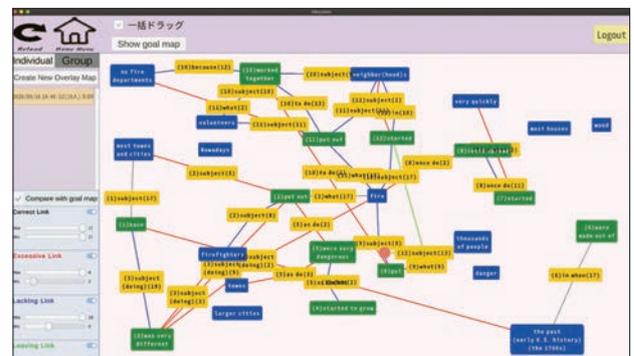


図5 組み立てられた概念マップの例

\*3 概念マップの長所を生かしつつ、作成した概念マップを容易に利用できるようにする構築手法。  
\*4 分析や意思決定に使われる「人の知見」とデータを融合させて要因分析が行えるモデリング手法。

解した英文についての概念マップを、提供した部品から学習者に組み立てさせた。部品は、教員が読解対象の英文に対して作成した概念マップを分解することで得られたものであり、元の概念マップと比較することで学習者が再構成した概念マップが自動診断できる。

ソフトウェアについては、一つの高校で、209 時限・人の実践利用を行い、運用可能性とデータ収集可能性を確認している。

## (2) 統計的 AI を用いた思考パターンの検出

思考経験データを学習支援に用いるために、「記号的 AI に基づく思考経験のデザイン」で得られる具体的な思考経験データから、思考パターンを検出する。自分が持っている情報を組み合わせて、妥当な構造を見つけることが思考であるとする、思考パターンには「構造（知識）」と「プロセス（知識の使い方）」の 2 種類があると考えられる。構造のパターンで組み立てた構造の多様性、プロセスのパターンで構造を組み立てていく手順の多様性を示すことで、どの部分からどのような順序で組み立てていくのかといったことや、どの部分で行き詰まるかといったことを分類できる。

また、思考パターンの検出には二つの意味がある。一つは事前に検討できた思考パターンが実際に出現するかの検証であり、もう一つは事前に検討できた以外のパターンの検出となる。このため、実際の思考経験データから統計的に思考パターンを検出する。また、検出した思考パターンを脳波データと結びつけ、脳の機能で説明することで、検出した思考パターンの妥当性を検証する。

実験・データの収集は広島大学で実施し、統計的 AI を用いたデータの処理・パターン抽出は広島工業大学に委託して実施。得られたパターンの解釈および意味付けは、広島大学と広島工業大学において協同で実施した。

## (3) 結論

サブテーマ「①記号的 AI に基づく思考経験のデザイン」に関しては、初年度に複数の実践参加教育機関を選定し、第 2 年度において教育現場と調整を行いながらシステム開発とテスト的利用を行った。第 3 年度に実践運用とデータ収集を行った結果、小中高校において 2,924 時限・人の実践を実施し、統計的 AI を用いた思考パターンの検出において必要十分となる 98,740 解答、368,650 ステップのデータが収集でき、本研究開発のアプローチが実践的に有用であることが確認でき

た。

サブテーマ「②実践を通して得られたデータに基づく統計的 AI を用いた思考パターンの検出」を実施する、データ収集が可能であることも実証できた。初年度に統計的 AI 手法の調査を行い、第 2 年度において過去のデータを用いた分析と視線や脳波などの生体データとの照合を含めて試み、ベイジアンネットワーク\*4 を用いたパターン検出が有望であると判断した。第 3 年度においては、ベイジアンネットワークを用いて収集したデータからのパターン分析を試み、正解の到達度を有意に向上させる状態および有意に下降させる状態の検出に成功した。②のテーマに関しては現時点で進行中の研究開発となるが、ここまでの成果によって、新規性のあるテラーメイドを実現する上での基礎となる思考パターンの抽出が可能であることが示せた。

サブテーマ「③思考パターンに基づくテラーメイド学習支援の開発」は実践データからのパターン抽出が可能となった第 3 年度に開始し、協力企業と合同でテーマ②で抽出したパターンに基づく学習者への適応的フィードバックを設計した。この適応的フィードバックは、従来の解答の正誤や誤答の種類に基づく適応的フィードバックに対して新規性があり、本研究開発の新規性を示せたと考えている。

サブテーマ「④ポータブルでセキュアな分散型データ管理」については、初年度および第 2 年度において協力企業と協議して学習ログの仕様を確定することができた。

## 4 まとめと今後の展望

サブテーマ①から③については、本研究開発を通して実証できたと判断している。また、これらの試みを一般化して実施するための基盤としての学習ログの仕様も確定できた。これらのことから、本研究開発では第 3 年度までに設定した、必要なデータの収集および必要な機能の設計・開発について、おおむね達成したと判断している。本研究開発の対象とした算数文章題および英文読解に関しては、教育現場での実践とその効果の実証を継続的に進めていくことで、思考経験を指向したテラーメイドの学習支援の社会実装事例としていく。