

「松下幸之助」 再現 AI システムの開発

大西 直¹ 高岸 智² 鎌田 理久¹ 山西 宏平² 神崎 雄介¹
小林 拓¹ 徐 曉琳¹ 菅原 陸¹ 杉浦 いぶき¹ 篠崎 友悠¹ 河村 岳²
¹ 株式会社松尾研究所

² パナソニックホールディングス株式会社

{t.onishi, s.xu}@matsuo-institute.com

{takagishi.satoru, yamanishi.kouhei, kawamura.takash}@jp.panasonic.com

{kamada.riku.am, yusuke.kanz, tkoba4932, murisu141421356,
ibukisugiura0403, dankuma43}@gmail.com

概要

本研究では、松下幸之助の思想や発言を忠実に再現する「松下幸之助」再現 AI システムを開発した。音声認識、返答生成、音声合成、動画生成の4つの AI モデルを統合し、ユーザとの自然な対話を可能にした。本システムは、膨大な著作物や音声データを学習対象とし、データクレンジングやリアルタイム処理の最適化、さらにドメイン知識の活用により、精度と自然さを大幅に向上させている。実験結果では、提案手法が返答生成の正確性、松下幸之助らしさ、応答の適切な長さなどの評価項目で高い性能を示し、史実に基づく自然な応答生成を実現した。これにより、企業理念の継承や文化的遺産の保存など、多様な領域での生成 AI 技術の新たな応用可能性が示唆された。

1 はじめに

松下幸之助 (1894–1989) は、松下電器産業株式会社 (現・パナソニックホールディングス株式会社) の創業者として知られ、その経営理念や思想は、同社をはじめ多くの企業経営者やビジネスパーソンに多大な影響を与えてきた。しかし、松下幸之助から直接薫陶を受けた人物が年々減少する中、その思想や経営理念を次世代へ正確に伝承することは容易ではない。

従来は主に書籍や講義型研修などによって思想や経営理念が伝承されてきたが、これらの手法は一般的に情報提供型であり、受講者が積極的に関与しづらい側面がある。そのため、これらの手段のみでは理念の真意を深く理解し、日々の実務に柔軟に活かすことが難しい場合がある。

本研究では、生成 AI 技術を活用し、松下幸之助の思想と話し方を再現する「松下幸之助」再現 AI システムの構築を目指す。本システムは松下幸之助の著作、講演記録、対談資料、映像、音声データを広範に学習することで、単なる過去の理念の再現にとどまらず、現代の課題に応じた動的な対話を実現する。これにより、より実践的な理解と継承が可能になると考える。

具体的には、音声認識、返答生成、音声合成、動画生成という4つの AI モデルを統合し、ユーザの音声入力に対して、あたかも松下幸之助本人が応答しているかのような自然な動画を生成する。こうしたアプローチは、企業理念の継承に限らず、教育や文化遺産の保存など、多様な領域での活用可能性を有するといえる。

2 関連研究

動画型 AI エージェントは近年、多様な分野で実用化が進んでいる。たとえば、HeyGen や Synthesia といったサービスは、現存人物の音声や動画をもとに高品質な合成映像を生成し、動画広告、教育、カスタマーサポートなど幅広い用途で利用されている [1, 2]。これらは現存人物を対象とするため、豊富なデータを活用しやすく、人物像や話し方、声質を忠実に再現できる。

一方、故人の再現では、利用可能なデータが限られ、不明瞭な音声や映像、古い表現方法などが大きな課題となる。また、当時の表現を現代的な文脈に適応する必要もあり、高度な自然言語処理技術が求められる。このような課題を克服するためには、データ統合のみならず、ドメイン知識の活用や徹底した品質管理が求められる。

本研究は、松下幸之助という特定のドメインに特化し、音声認識、返答生成、音声合成、動画生成を統合することで、史実に基づく自然な対話を実現する点で既存技術とは一線を画すものである。

3 提案手法

本研究が提案する「松下幸之助」再現 AI システムは、図 1 に示す 4 つの AI モデルから構成される。ユーザの音声入力を受け音声認識、返答生成、音声合成、動画生成を行う。本システムの精度を向上させるため、膨大なデータに対し以下の 3 点を重点的に実施した：

1. データクレンジング
2. 処理時間の高速化
3. ドメイン知識による改善

3.1 データクレンジング

本システムの精度向上において、データクレンジングは不可欠なステップである。株式会社 PHP 研究所より受領した著作物や講演記録に含まれる古い言い回しの修正や、ノイズを含む音声データの前処理を行った。これにより、返答生成や音声合成の精度を高め、自然な対話を実現できる。

3.1.1 テキスト前処理

本研究における返答生成モデルの初期段階では、回答内容が短く抽象的であり、具体的な実体験や思想的背景が欠如しているとの指摘も受けた。この課題を解決するため、以下の改善プロセスを実施した。

1. データベース設計の最適化

松下幸之助の発言集をもとに、QA データベースを構築した。具体的には、以下の手法を採用した：

- 標準語化と重複除去

発言集の古い表現を標準語に修正し、意味が重複する発言を排除することで、検索効率と品質を向上。

- 質問データの生成

過去の発言集をもとに、その内容が回答になりうるような具体的かつ自然な質問を生成。曖昧または不自然な質問は除外し、発言 1 つに対し質問 1 つが紐づく形式でデータを整理。

2. 参照文書数の増加

短く抽象的な回答に留まらず、回答に多様な具体

性を付加するため、返答生成時に参照する文書数を 5 個から 10 個に増加した。その結果、実体験や哲学的背景を含んだ回答生成を実現した。

3. モデル構成の検討と選定

文書検索には RAG (Retrieval-Augmented Generation) [3] 手法を採用した。当初、異なる検索手法の利用やそれらの併用も検討したが、適切に設計されたデータベースによって十分な精度が達成可能であることと、また返答速度が最も早くなることを考慮し、Naive な RAG を採用した。

3.1.2 音声前処理

対象とした音声データは、過去に録音されたものであるため、ノイズが多く含まれていた。このため、音声データに対して以下の前処理を実施した：

- 背景雑音や BGM の除去
- 発話内容と文字起こし結果の不一致を修正
- 無音区間の自動除去
- 松下幸之助以外の音声や不明瞭な音声を排除

これらの処理を通じて、音声合成の学習に用いるための高品質なデータセットを構築した。

3.2 処理時間の高速化

動画生成は処理フローにおける主要な要素であり、入力となる合成音声の長さが増加するにつれて動画生成モデルの推論時間が線形的に増加するという課題が存在する。本課題を解決するために、音声分割手法を活用して合成音声を動的に分割し、各セグメントに対して動画生成モデルを適用する手法を採用した (図 2)。

提案手法では、分割された音声セグメントごとに逐次的な動画生成を実現し、ユーザが生成済みの動画を視聴している間に残りの動画生成プロセスを並行して進行させることが可能である。この手法により、全ての動画が生成完了するまで待機する必要がなくなり、初回の音声セグメントに対応する推論時間のみを要する効率的な処理を実現した。

本研究では、松下幸之助の時代背景を考慮し、高画質や高 FPS は必須ではないとの判断から、動画生成における高画質化プロセスを省略した。また、動画の FPS を従来の 25 から 15 に低下させる最適化を実施した。この軽量化により、27 秒の合成音声に対する動画生成の推論時間は従来の 27.37 秒から 18.66 秒へと約 32 % の削減を達成した。これにより、リ

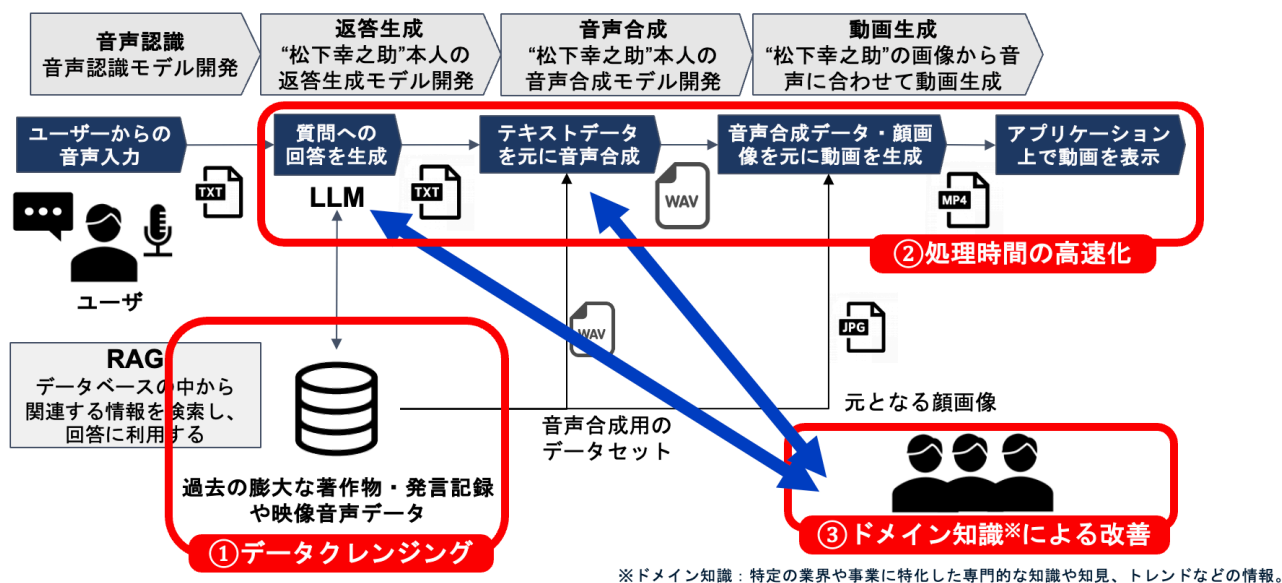


図1 開発システム全体構成

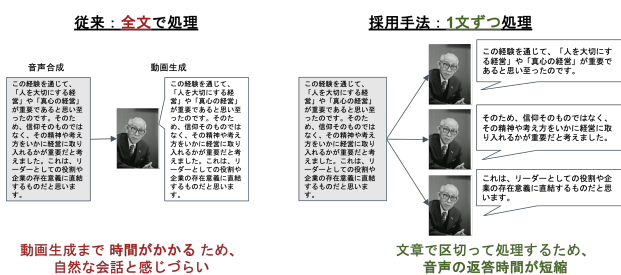


図2 処理時間の高速化システムの構成

ソース効率が向上し、リアルタイム性の更なる強化が可能となった。

3.3 ドメイン知識による改善

松下幸之助の特徴や発言意図をより忠実に反映した回答を生成することを目的とし返答生成モデルの入力プロンプトの再設計を行った。この再設計では、以下の2つの改善を中心に取り組んだ。

- 松下幸之助の発言が行われた時代背景や社会的状況

具体的には、戦後の経済復興期や世界恐慌の影響下など、発言の文脈を形成する重要な出来事や状況をモデルに与える形でプロンプトを設計した。これにより、生成される回答が単なる言語的模倣に留まらず、発言の背景を考慮した自然で説得力のあるものとなった。

- 松下幸之助の重要な思想や哲学、過去の代表的な発言を基にプロンプトを調整

特に、「経営理念」や「人間尊重」といった思

想に紐づく具体的な発言を優先的にモデルに参照させる仕組みを採用した。これにより、生成される回答が松下幸之助の核心的な思想を反映しつつ、質問の意図に即したものとなるよう改善された。

生成された返答や音声については、松下幸之助研究者や彼から薫陶を受けた株式会社 PHP 研究所とパナソニックオペレーションズエクセレンス株式会社ブランド・コミュニケーション部門ブランド戦略センター歴史文化コミュニケーション室（歴史文化室）全8名の合計3回の点数評価とフィードバックをもとに繰り返し改善を行った。評価軸は以下の通りである。

- 返答生成 正確さ：史実に基づいた回答が生成されているか
- 返答生成 松下幸之助らしさ：松下幸之助の特徴的な言葉遣いが含まれているか
- 返答生成 応答の長さ：質問内容に適した長さであるか
- 音声合成 松下幸之助らしさ：声質や話し方が実際の松下幸之助に近い
- 音声合成 話し方のスピード：話すスピードが適切であるか

4 結果

本研究の改善前システムは、収集可能な松下幸之助関連テキスト・音声データをフィードバックサイ

クルを導入せずに構築したベースラインモデルであった。この改善前モデルと比較することで、本研究で導入したデータクレンジング、ドメイン知識の導入などの改善点の有効性を評価する。

4.1 専門家による評価結果

本研究では、返答生成および音声合成に関する5つの評価項目について改善前後における性能評価を実施し、各評価項目について数値的な比較を行った。その結果を表1に示す。これらのスコアは、3.3章で述べた全8名からの5段階評価（1が最低、5が最高）の平均点である。

表1 評価項目の改善前後の比較

評価項目	改善前	改善後
返答生成_正確さ	2.70	3.90
返答生成_松下幸之助らしさ	3.06	4.45
返答生成_応答の長さ	3.51	4.60
音声合成_松下幸之助らしさ	2.79	3.58
音声合成_話し方のスピード	3.35	3.96

4.2 改善前後の返答生成結果

システムの改善前後で、生成されるテキストの品質向上が確認された。例として、「経営で一番重要なことは何でしょうか？」という質問に対する回答を比較する。

質問: 経営で一番重要なことは何でしょうか？

改善前の出力:

経営において一番重要なことは、やはり「**熱意**」であります。熱意がなければ、どんなに知識や経験があっても成功は難しいと思うんです。熱意をもって、全力で取り組むことが肝要であります。

改善後の出力:

経営において一番重要なことは、やはり「**人間尊重**」の精神であります。経営というのは、単に利益を追求するだけではなく、人々の生活を豊かにし、社会全体の繁栄に貢献することが求められるんです。そのためには、まず従業員一人ひとりを大切に、その能力を最大限に引き出すことが肝心であります。

また、**経営理念**をしっかりと持ち、それを実践することも非常に重要であります。

改良後のモデルは、初期段階と比較して「松下幸之助らしさ」が顕著に向上した。例えば、質問「経

営で一番重要なことは何でしょうか？」に対して、「経営理念の確立」という要素が自然に回答に組み込まれるようになった。さらに、関連する実体験や哲学的背景が回答に付加されることで、研究者から「松下幸之助の思想を忠実に再現している」と高く評価された。

これらの結果から、本研究で導入したデータクレンジング、ドメイン知識の導入が返答生成および音声合成の両面でシステムの性能を向上させる有効なアプローチであることを示している。

5 結論

本研究では、松下幸之助の発言や思想を忠実に再現する「松下幸之助」再現 AI システムを開発した。このシステムは、音声認識、返答生成、音声合成、動画生成の4つの AI モデルを統合し、ユーザとの自然な対話を実現することを目指している。本研究では、データクレンジング、リアルタイム処理の最適化、ドメイン知識を活用した改善によって、返答生成の正確性、応答の適切な長さ、そして音声や話し方の自然さといった重要な評価項目で顕著な改善を示した。特に、RAG を活用しドメイン固有のデータセットを適用することで、史実に基づいた回答生成が可能となり、より松下幸之助らしい応答が得られた。

本研究における返答生成の改善は、データベース設計とプロンプト設計を中心に行った結果、複雑なアルゴリズムの追加なしに、検索精度および生成精度の向上を実現した点に特徴がある。このアプローチは、類似の再現 AI 開発においても応用可能であると考えられる。

参考文献

- [1] HeyGen, "AI Video Generator," HeyGen, 2023. <https://www.heygen.com/>
- [2] Synthesia, "About Synthesia," Synthesia, 2023. <https://www.synthesia.io/>
- [3] Lewis, P., et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," Advances in Neural Information Processing Systems (NeurIPS), 2020.