

対話型AI「YourSelf LM」による自己理解促進支援システムの開発

作成日付: 2025年8月9日

作成者: Nikeri

1. 概要・要旨

情報通信技術（ICT）によって全世界が接続され、地域による情報格差が緩和される一方、文明社会はその急速な発展に伴う新たな摩擦に直面している。開発者のいない地域におけるデジタル・ディバイド（情報格差）や、ユーザーの関心を惹きつけることで収益を最大化するSNSのビジネスモデルがもたらすコミュニケーションの質の変化は、特に若年層において深刻な社会問題となりつつある。

本稿ではこれらの問題の中でも、デジタル技術がもたらす利便性と効率性の裏側で進行する「対話機会の減少」と、それに伴う「自己解像度の低下」という課題に焦点を当てる。自己を深く省みる機会が失われることは、自己肯定感の低下や、人生の重要な局面における意思決定の困難に直結しうる。

本プロジェクトの目的は、この課題を解決するため、対話型AI「YourSelf LM」を開発し、ユーザー自身が感情、思考、行動の背後にある根本的な原因・理由を理解する「自己理解」を促進することである。本システムは、ユーザーが提供するテキストデータから、その人固有の価値観や文体の「癖」を学習し、AIアバターとして再現する。ユーザーはこのAIアバターとの対話を通じて、自身を客観的に見つめ直す機会を得る。

このアプローチにより、若年層の健全な自己形成を支援するとともに、キャリア設計をはじめとする人生の重要な意思決定の質を向上させることを目指す。

2. 背景

2.1 デジタル時代のコミュニケーションと自己認識

現代社会は、常時接続が当たり前の環境となった。これにより、コミュニケーションの形態は大きく変容し、即時性や簡潔さが重視される傾向にある。この変化は、他者との深い対話や、自己の内面と静かに向き合う時間を減少させた。自己とは、他者との関わりの中でその

輪郭を認識していく側面を持つため、こうした対話機会の減少は、自身が何を考え、何を感じているのかを把握する「自己解像度」を低下させる一因となっている。

2.2 既存の自己分析ツールの限界

これまでも、自己理解を助けるためのツールとして、各種の性格診断テストや心理検査が存在した。これらは一定の有効性を持つものの、多くは以下の限界を抱えている。

- **一過性:** 診断は一度きりで、その後のユーザーの変化に追従しない。
- **対話性の欠如:** 静的な結果を提示するのみで、対話を通じて疑問を解消したり、結果を深掘りしたりすることができない。
- **画一性:** すべてのユーザーに同じ質問形式を適用するため、個々の文脈や特性を十分に反映できない。

これらの限界は、ユーザーが継続的に自己と向き合い、その動的な変化を捉える上での障壁となっている。したがって、パーソナライズされ、対話的かつ継続的な自己探求を可能にする新たな支援手法が求められている。

3. 実装

本システム「YourSelf LM」は、プライバシーを最優先する**基礎設計**と、リアルタイム性と分析の深層化を実現する**発展設計**を融合させることで、前述の課題解決を目指す。

3.1 システムアーキテクチャの基本思想

本システムの核心は、ユーザーデータの所有権をユーザー自身に留めつつ、高度なAIとの対話体験を提供することにある。これを実現するため、データ管理の思想は初期構想から一貫させ、その上で動作パフォーマンスを飛躍的に向上させる技術を導入する。

3.2 データ管理とプライバシー保護（基礎設計）

ユーザーのプライバシー保護は、本システムの根幹をなす設計原則である。

1. **データ入力と初期分析:** ユーザーは自身のチャット履歴や文章などを提供する。これらのデータは、まず自然言語処理ライブラリGiNZAにより、語彙、文体、構文といった基本的な言語的「癖」が分析される。
2. **ユーザー主権のデータストレージ:** 分析された人格データや対話ログの原本は、運営側のサーバーではなく、ユーザー個人のGoogle Driveアカウント内の専用フォルダ

に保存される。これにより、ユーザーは自身のデータに対する完全なコントロール権を保持する。

3.3 高速・深層対話の実現（発展設計）

基礎設計のままでは、分析に長時間を要し、コストも高くなる。この課題を以下の技術的アプローチで解決する。

1. 処理の高速化:

- **非同期・増分処理:** 初回の人格分析はバックグラウンドで非同期に実行する。2回目以降は、新たに追加された対話データのみを学習する**増分学習**により、人格更新にかかる時間を従来の48分から数秒へと短縮する。
- **高速RAGシステム:** ユーザーの心理的要素や価値観を格納するデータベースには、高速な**インメモリDB (Redis) **を採用。これにより、対話の文脈に応じた関連情報を10ms台で検索し、応答生成の高速化に貢献する。
- **推論の最適化:** 推論処理にはTensorRTやFlashAttention-2のような高速推論カーネルを導入し、AIの応答速度を最大化する。

2. 低コスト・高品質な応答生成:

- **高性能軽量モデル:** ベースモデルにはPhi-3-miniのような軽量な言語モデル (SLM) を採用する。
- **知識蒸留:** 高性能な大規模モデルの思考プロセスのみを軽量モデルに学習させるChain-of-Thought Distillationという手法を用いる。これにより、APIコストを抑えつつ、応答の論理的整合性や説明能力を維持する。
- **効率的ファインチューニング:** 4-bit QLoRAという技術を用い、少ない計算リソースで効率的にモデルのチューニングを行う。

3. 分析の深層化:

- **心理学的フレームワークの統合:** Big Five理論だけでなく、ナラティブセラピーといった心理学的知見を統合する。これにより、ユーザーの内面をより多角的かつ深く分析し、豊かなフィードバックを生成することが可能となる。

4. 期待される効果

4.1 技術的効果

- **処理速度の劇的向上:** 2回目以降の応答生成を数秒レベルに短縮し、リアルタイム対話を実現する。
- **運用コストの削減:** 軽量モデルと知識蒸留の活用により、APIコールなどの運用コストを最小化する。
- **説明可能性の向上:** 軽量モデル単独で応答理由を説明できるアーキテクチャにより、AIの透明性を確保する。

4.2 ユーザー体験の向上

- **リアルタイム対話:** ストレスのない自然な対話を通じて、没入感のある自己探求体験を提供する。
- **パーソナライズされた支援:** ユーザーの内面変化に動的に追従し、常にその人らしい応答を返すことで、継続的な自己理解を支援する。
- **深層理解の促進:** 多角的な心理学的視点からのフィードバックにより、ユーザー自身も気づけなかった自己の側面に光を当てる。

4.3 社会的インパクト

- **精神的ウェルビーイングへの貢献:** 自己理解を深めることは、自己受容や精神的な安定に繋がる。本システムは、若年層のメンタルヘルス向上に寄与する可能性がある。
- **意思決定の質の向上:** 進路選択やキャリア設計など、人生の重要な局面において、自己の価値観に基づいた、より納得感のある意思決定を支援する。

5. 現状と課題

5.1 現在の開発状況

本プロジェクトは、コンセプト実証（Proof of Concept）を終えた段階にある。基礎的なプロトタイプを用いた実験により、Big Five理論に基づきユーザーの基本的な人格を再現すること、および重み付きタグ検索によって文脈に応じた応答を生成することの基本的な有効性は実証済みである。現在は、本提案書で示した発展的設計に基づき、実用的なシステムへと昇華させるフェーズを目指している。

5.2 今後の検証課題

本提案の実現には、以下の技術的・心理学的仮説の検証が不可欠である。

- **技術的課題:**

1. **増分学習の精度維持:** 差分データのための学習で、人格の一貫性を損なうことなく、どの程度の精度を維持できるか。
2. **知識蒸留の効果:** 思考プロセスを蒸留する手法で、どの程度心理的な説得力や説明品質を軽量モデルに移植できるか。
3. **合成データの品質:** AIを用いて生成した学習データが、現実の対話データと同等の品質を持ち、バイアスを増幅させないか。

- **倫理的・心理学的課題:**

1. **心理的複雑性の単純化:** 複数の心理モデルを統合してもなお、人間の複雑な内面を単純化しすぎるリスクはないか。
2. **AIへの過度な依存:** ユーザーがAIの分析に依存し、自律的な思考を放棄するリスクにどう対処すべきか。
3. **支援の境界線:** 「自己理解支援」と「心理療法」の境界を明確にし、ユーザーに誤解を与えないための設計が求められる。

6. まとめ

本稿では、現代社会における対話機会の減少と自己解像度の低下という課題に対し、対話型AI「YourSelf LM」を開発することで、新たな自己理解支援の形を提案した。プライバシーを保護する基礎設計の上に、高速化と深層化を実現する発展設計を組み合わせることで、安全かつリアルタイムな自己探求体験の提供を目指す。

基礎的な有効性は確認済みである一方、その実用化に向けては、本稿で述べたような技術的・倫理的な課題が残されている。これらの課題を解決し、本プロジェクトの社会的意義を最大化するためには、多様な分野の専門家からの客観的な評価と建設的な助言が不可欠である。

つきましては、各分野の専門家の皆様に本提案をレビューいただき、その妥当性評価と、今後の開発方針に関するご指導を賜りたく、ここに提案する次第である。