

# 発明の名称: AI を活用した多次元感情・認知統合メンタルケアシステム

## 技術分野

本発明は、人工知能およびメンタルヘルスケアの分野に関するものである。特に、対話型 AI を用いて、ユーザーの感情・認知状態および生体データを統合的に分析し、メンタルケアに応用する技術に関する。

## 背景技術

メンタルヘルスチャットボットや仮想セラピストは、カウンセリングや認知行動療法 (CBT) の手法でユーザーを支援するツールとして知られている。しかし従来技術には以下のような制限がある。

- **感情理解の限界:** 一般的なチャットボットは単純な感情分析や定型スクリプトに頼り、複雑な感情状態を十分に捉えられない。人間の感情は多次元的な性質を持つが、従来システムでは例えば不安と悲しみの同時存在などを検出できず、ユーザーの微妙な気分変化を見逃しがちである。
- **認知の歪み検出の欠如:** ユーザーの発話に含まれる認知の歪み (例: 過度の一般化や白黒思考、「自分は全てに失敗する」など) はメンタルヘルス不調の一因となる。既存のチャットボットは、自然な会話の中からこれらの歪みを的確に認識し、それに対処するような対話を行うことができない。その結果、ユーザーに対し一般論や直接的な忠告しかできず、かえって抵抗感や無力感を与えることがある。
- **生体情報の未活用:** スマートウォッチ等のウェアラブルデバイスは心拍数や睡眠など精神状態に関連する生体データを取得できるが、従来のメンタルヘルス AI はこれら客観的な指標を考慮していない。したがってユーザー状態の評価において重要な文脈 (例えば心拍上昇による隠れた不安、睡眠不足による脆弱性など) が見落とされていた。
- **単調なインターフェース:** 多くの既存システムはテキストまたは単純な音声応答のみで、人間らしい視覚的フィードバックを欠く。一部にアバターを用いる例もあるが表情が乏しいか定型的であり、共感や信頼感を十分に与えられな

い。また音声出力においても、感情内容と声のトーンが合致しない場合があり、治療効果を高めるには不十分であった。

- **外部サービスへの依存:** 多くの AI 対話システムはサードパーティの大規模言語モデルやクラウド API に依存している。このため機密性の高いメンタルヘルスデータのプライバシーリスクや、対話内容・治療戦略を開発者が細かく制御できないという問題がある。また外部 API 中心では、音声トーンの調整やセンサーデータ統合など専用アーキテクチャで可能なリアルタイム最適化が困難である。

## 本発明が解決しようとする課題

本発明は上述の課題を解決するため、ユーザーの感情状態を多次元で精細にモデル化し、会話中に表出される認知の歪みを検知して適切に対処し、生体センサーデータも考慮した応答適応を行い、加えて感情に寄り添ったアバター表現によるフィードバックを提供する統合的な AI メンタルケアシステムを提供することを目的とする。具体的には、本発明は以下の点に着目している:

1. **多次元感情モデルによる正確な感情把握:** ユーザーの複雑な感情を原子価（ポジティブ～ネガティブ）、覚醒度（活発～沈静）など複数軸のベクトルとして表現することで、一面的ではない豊かな感情理解を可能にする。時間経過による感情ベクトルの変化を追跡し、感情バランス（均衡状態）への動きを評価できるようにする。
2. **治療的対話による認知再構成:** 共感的にユーザーの話を受容しつつ、その中に含まれる否定的思考パターンや非合理的な信念（認知の歪み）を検出する。そしてそれを直接指摘するのではなく、質問や提案を通じてユーザー自身が気づきを得て思考を再構築できるような認知行動療法的対話を実現する。これによりユーザーが防御的にならずに、より健全な認知パターンを身につけられるよう支援する。
3. **生理的信号の活用:** ウェアラブル等から得られる心拍、血圧、皮膚電気反応、睡眠の質などの生体情報をリアルタイムに対話へ統合し、ユーザーの報告と客観指標との差異を検知して洞察を深める。例えばユーザーが「大丈夫」と言っているにもかかわらず心拍上昇が見られる際には潜在的ストレスを考慮し、逆に主観的には不安でも生体的には安定していれば安心を促すなど、より精度の高いケアにつなげる。
4. **没入感のあるフィードバック:** VRM 形式の 3D アバターを用い、対話内容に応じて表情や身振りを変化させることで対面で対話しているような臨場感を提供

する。また、出力音声もリアルタイム音声変換(RVC)技術により、応答の感情に沿った声色・トーンに変換する。ユーザーが落ち着くよう穏やかな声や、励ましとなる明るい声など、感情にマッチした音声出力で共感と安心感を高める。

5. **独立型かつプライバシー保護:** コアとなる対話処理を外部に依存しないアーキテクチャとし、ユーザーデータは必要最小限だけクラウド送信し、大部分を端末内で処理する。これによりネットワーク遅延を抑えつつ、個人情報を保護し、安全性の高い環境(オフラインや閉域ネットワーク)でも動作可能とする。独自開発の AI モデルにより、治療対話に最適化された微調整やアップデートを容易にし、外部 API では実現困難な統合機能を提供する。

## 発明の概要

本発明に係る AI メンタルケアシステムは、新規の構成要素を組み合わせることでユーザーの心理状態を総合的にモニタリングし、リアルタイムで治療的介入を行うものである。一実施形態において、本システムは以下を含む:

- **多次元感情分析モジュール:** ユーザーの入力(テキスト、音声トーン、表情データなど)および生体センサーデータを解析し、前記した感情空間内でユーザーの感情状態を示す多次元ベクトルを算出するモジュール。このベクトルは対話のたびに更新され、ユーザーの感情履歴や傾向を定量的に表現する。
- **認知歪み検出モジュール:** ユーザーの発話テキストを自然言語処理によって分析し、認知の歪み(例:白黒思考、破局化、自己批判など)に該当するフレーズやパターンを識別するモジュール。直接の指摘は避けつつ、内部的にフラグを立てて対話方針に反映させる。
- **治療対話生成モジュール:** 前記感情ベクトルと認知歪みの情報を基に、心理療法の知見に沿った応答を生成するモジュール。具体的には、共感を示す表現と、必要に応じてユーザーの認知を優しく再構築する質問や提案を組み合わせた対話フローを組み立てる。この際、大規模言語モデル(LLM)を利用して自然な文章を生成しつつも、あらかじめ定義した治療戦略(認知行動療法に基づくスクリプト)に沿うよう統制する。
- **アバター出力モジュール:** 前記生成された応答をユーザーに伝達するために、VRM アバターキャラクターを用いた表示と、感情一致音声(Emotion-Matched Voice)での音声出力を行うモジュール。アバターは対話内容に応じた表情変化やボディランゲージを示し、音声は応答内容の感情トーンに合うようリアルタイムに変換される。

- **生体情報連携モジュール:** ウェアラブル等のセンサーデバイスから心拍数や睡眠データ等を安全に取得し、前記多次元感情分析に組み込むモジュール。センサーデータはユーザー許可のもとプライバシーに配慮して取り扱われ、対話中のユーザー状態評価や異常検知(例:ストレスレベルの急上昇など)に活用される。
- **クラウド・エッジ分散処理構成:** システムはユーザーデバイス側のクライアントアプリと、クラウドまたはローカルサーバ側の AI プラットフォームに分散して実装される。クライアント側で音声認識やアバター描画などリアルタイム性とプライバシー重視の処理を行い、サーバ側で感情分析や対話生成など高負荷の AI 処理を行う構成とする。両者は暗号化通信で連携し、ユーザー体験の滑らかさとデータ保護を両立する。
- **セキュリティおよびパーソナライズ:** システム全体にセキュリティ対策が組み込まれ、ユーザーデータや AI モデルが不正に改ざん・複製されることを防止する。例えば通信やモデルファイルの暗号化、認証キーによるモジュール認証などにより、許可のないシステム改変や外部からの悪用を阻止する。また各ユーザーの対話履歴や嗜好に基づき応答を調整するパーソナライズ機能も備え、プライバシーを確保しつつ個別ニーズに応じた対応を可能にしている。

## 図面の簡単な説明

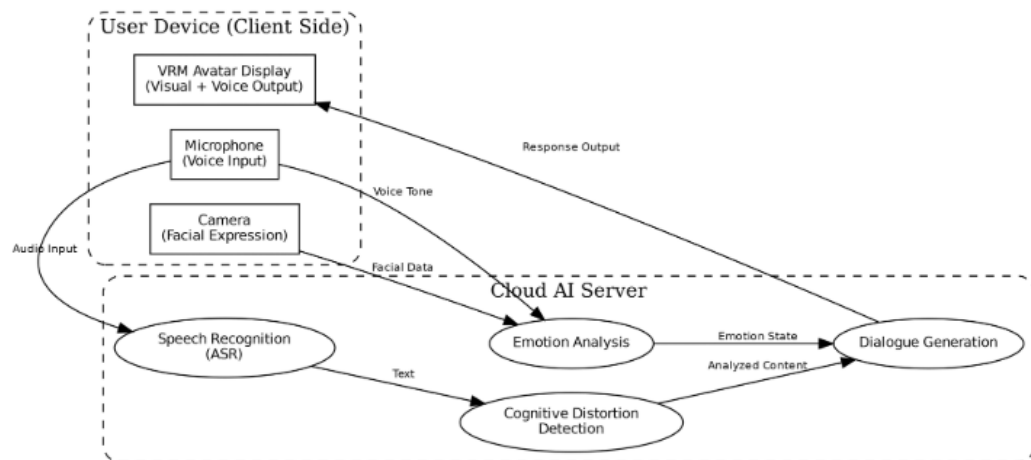
図 1 は、本発明の一実施形態によるシステム構成(システムアーキテクチャ)の概略図である。

図 2 は、ユーザー入力からアバター応答生成までの対話処理フローの一例を示すフローチャートである。

図 3 は、ユーザー発話の例とそれに対応して検出される認知の歪みカテゴリを示す概念図である。

## 実施形態の詳細な説明

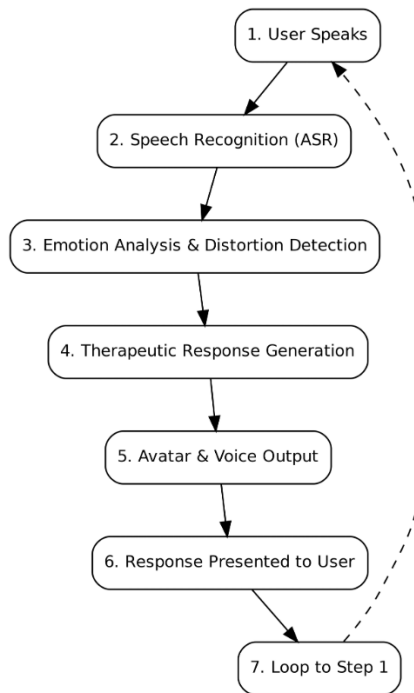
## システムアーキテクチャの概要



システムのクライアント・サーバ構成を示す概略図。クライアントデバイス(左)で音声・映像入力とアバター出力を行い、クラウド AI サーバ(右)で感情・認知分析と応答生成を担当する。(図 1: System Architecture Overview)

本システムは、ユーザー側クライアントデバイスとクラウド/サーバ側プラットフォームに大別される。図 1 に示すように、クライアント側では音声入力やカメラ映像、ウェアラブルセンサからのデータ取得を行い、サーバ側で感情・認知分析および対話生成の核心処理を行う。ユーザーはマイクとカメラを備えたデバイス(スマートフォン等)を用いてシステムと対話し、そのデバイス上で VRM アバターによる応答表示を受ける。センサーデバイスから得られる生体データもクライアントデバイス経由でサーバに送信され、AI 解析に統合される。クラウドサーバ上では、音声認識、表情解析、感情ベクトル計算、認知歪み検出、応答生成といった各モジュールが密接に連携する。処理の流れとしては、ユーザー発話がクライアント側でテキスト化・前処理された後サーバに送られ、解析結果に基づく応答が生成されて再びクライアント側に送信され、アバターと音声によってユーザーに提示される。なお、通信は全て暗号化され、プライバシーが確保される。

## インタラクションの処理フロー



対話処理フローの一例を示す図。ユーザー発話の入力から、音声認識、感情ベクトル算出、認知歪み検出、応答生成を経て、アバターと音声でユーザーにフィードバックするまでの流れを示す。(図 2: Therapeutic Interaction Flow)

図 2 に基づき、本発明の対話処理フローの一例を説明する。まずユーザーがマイクに向かって感情を含む発話を行うと(ステップ 1)、クライアント側で音声認識が実行されテキスト化される。同時にカメラ映像や音声から表情や声のトーンといった感情の手がかりも抽出される(ステップ 2)。得られたテキストと特徴量はサーバ側に送られ、感情分析モジュールによって多次元感情ベクトルが計算される(ステップ 3)。さらに発話内容に対して認知歪み検出が行われ、例えば「もう何もかもうまくいかない」といった表現から「全か無か思考(白黒思考)」の歪みがあると判断される(ステップ 4)。これらの情報を踏まえ、対話生成モジュールが適切な応答を決定する(ステップ 5)。例えば、ユーザーの感情を受け止めつつ、歪んだ認知に対して間接的に視点を変える質問を組み込んだ応答を構築する。生成されたテキスト応答はクライアント側に送り返され、VRM アバターの表情と音声変換によってユーザーに提供される(ステップ 6)。ユーザーはアバターからの共感的な声掛けや表情を受け取り(ステップ 7)、対話が 1 サイクル完了する。このプロセスが繰り返されることで、ユーザーの感情状態の変化に応じた継続的な支援が行われる。

## 多次元感情ベクトルモデル

本システムが用いる感情モデルは、感情を複数の軸で表現する多次元ベクトルとして  
いる。典型的な軸として、快-不快(原子価)、覚醒度(高エネルギー-低エネルギー)、  
主導感(コントロール感-無力感)などが含まれる。各対話においてユーザーの発話  
内容、音声トーン、表情解析結果、生体指標を総合してこれら軸上のスコアを算出  
し、感情状態をベクトルとして表現する。例えば、ある瞬間のユーザー状態(原子価  
=-0.5, 覚醒度=+0.8, 主導感=-0.2,...)のように数値化する。このベクトルは時間ととも  
に軌跡を描き、ユーザーの感情変化を示す。システムは理想的なバランス状態(各軸  
が中間値付近)も定義しており、ユーザーの現在ベクトルとの差分から、より均衡の  
取れた方向へ誘導するよう対話戦略を調整する。

## 治療対話エンジンと認知歪みへの対処

対話生成の中核となるエンジンは、心理療法の知見に基づいたシナリオに沿って応  
答を組み立てるよう設計されている。システムはユーザーからの各発話に対し、更新  
された感情ベクトルと検出された認知の歪みフラグを考慮して、適切な治療的応答を  
決定する。例えば、ユーザーが「自分は何をやっても失敗する」と発言した場合、これ  
は「全か無か思考」の歪みであると内部で認識される。その際、対話エンジンは単に  
「そんなことはありません」と否定するのではなく、「最近何かうまくいったことはありません  
でしたか？」といった問いかけを返すことで、ユーザー自身に例外や可能性を考  
えさせるよう誘導する。このように、検出された歪みの種類に応じて複数用意された  
対話パターンの中から適切なものを選択し、ユーザーの視野を広げ認知を柔軟にす  
るサポートを行う。

応答生成には必要に応じ大規模言語モデル(GPT 等)が活用されるが、その出力は  
上記の治療対話フレームに沿うよう統制される。具体的には、システム独自の対話管  
理が LLM へのプロンプトや温度設定を調整し、常に共感的でユーザーに寄り添った  
トーンを維持しつつ、歪みへの介入を織り交ぜた発言を生成する。これにより高度な  
言語運用能力と治療指向性を両立させている。

## VRM アバターインターフェース

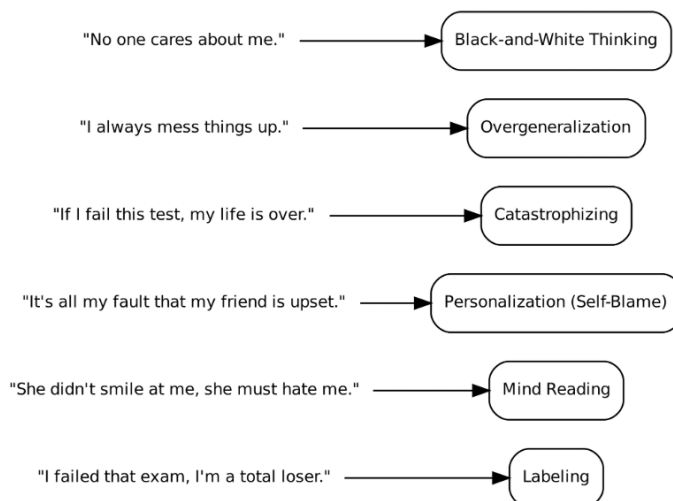
クライアント側のユーザーインターフェースには 3D アバター(VRM 形式)が用いられ  
る。アバターはユーザーに対する視覚的エージェントとして振る舞い、AI の応答内容  
に合わせて表情やジェスチャーを変化させる。例えば、ユーザーが落ち込んだ内容を  
話した場合、アバターは悲しそうな表情を見せる。逆に前向きな進展があれば微笑  
む。この表情変化はリアルタイムの対話において重要な非言語的フィードバックとな

り、ユーザーに安心感と共感を伝える。

また音声出力には Emotion-Matched Voice 機能が組み込まれている。テキストで生成された応答は、一旦ニュートラルな音声（例えば音声合成）に変換された後、RVC（Retrieval-based Voice Conversion）技術によって目標とする感情トーンに音声特性が変換される。これにより、例えば慰めが必要な場面では優しく穏やかな口調、励ます場面では明るく活力のある口調といったように、声の調子そのものもメッセージの感情に即した形でユーザーに届けることが可能となる。

アバターと感情一致音声による出力により、ユーザーは単なるテキスト以上に豊かな対話体験を得ることができ、AI に対する親近感や信頼感が高まる。これらの要素は特にメンタルケアにおいて、ユーザーが心を開き継続して対話に臨むための重要な要因となる。

## 認知歪みマッピングの例



ユーザー発話例と認知の歪みカテゴリの対応を示す図。否定的な自己表現がどの歪みに分類されるかを例示している。（図 3: Cognitive Distortion Mapping）

図 3 は、ユーザー発話の具体例とそれに対応する認知の歪みカテゴリーを示したものである。例えば「誰も私のことを気にしてくれない」という発言は「全か無か思考」に分類される。同様に「これは完全な災難だ」は「破局化」に、「全て自分のせいだ」は「過度の自己非難」にそれぞれ該当する。このように本システムでは、ユーザーの否定的な表現を既知の認知歪みパターンにマップし、内部的にタグ付けしている。対話エンジンはこのタグをトリガーとして適切な応答戦略（例えば全か無か思考には段階的に可能性を探る質問を返す等）を選択する。これによりユーザー自身がより現実的で柔軟な視点を得られるよう支援している。



## 倫理的配慮と悪用防止

本発明のシステムは、メンタルヘルス支援に特化したものであり、意図しない用途への利用を明確に排除している。具体的には、宗教的または政治的な勧誘、思想洗脳やマインドコントロールの手段、ユーザーの不健全な感情的依存を意図的に誘発するような使い方は、本システムの適用範囲から除外されている。このような不適切な対話要求が検出された場合、システムは応答を生成しないか、安全なガイドラインに従った対応のみを行うよう設計されている。

さらに、システムには構造的な悪用防止策が組み込まれている。対話データや AI モデルは高度に保護されており、外部からの改ざんや不正なクローン作成を防ぐセキュリティモジュールが含まれる。例えば、すべてのモデルファイルには電子署名が付され、真正性検証に失敗した場合には起動しないようになっている。また、出力されるコンテンツにもデジタル透かし等を挿入し、第三者がシステムを模倣または不正利用する試みを検知できるようにする。これらの対策により、本システムは提供者の意図した安全な範囲でのみ動作し、ユーザーの信頼を損なうような悪用が起こらないようになっている。

## 特許請求の範囲

【請求項1】 ユーザーのメンタルケアのための AI ベースの対話型システムであって、ユーザー入力データを解析して複数の感情次元にわたるユーザーの感情状態ベクトルを計算するよう構成された多次元感情分析モジュールと、ユーザーの発話内容を分析して認知の歪みを検出するよう構成された認知歪み検出モジュールと、前記感情状態ベクトルおよび検出された認知の歪みに基づいてユーザーに提供する治療的な対話応答を生成する対話生成モジュールと、前記対話応答に対応する表情を表示する VRM アバターおよび感情一致音声によって応答を出力するアバター出力モジュールと、ユーザーの生体情報を取得するウェアラブルセンサと、該センサからの生体データを前記多次元感情分析モジュールに供給する生体情報連携手段と、前記モジュール群をユーザーデバイス上のクライアントアプリケーションとクラウドまたはローカルサーバ上の AI プラットフォームとに分散配置し、ユーザーのプライバシーを保護しつつリアルタイム処理を実現する分散処理手段と、前記システムのアルゴリズムおよびデータに対する不正な改ざん、複製または悪用を検知・防止するセキュリティモジュールと、を備えることを特徴とする対話型 AI メンタルケアシステム。

【請求項2】 前記多次元感情分析モジュールは、ユーザーの感情状態を少なくとも原子価および覚醒度の二軸を含む感情空間内のベクトルとして算出し、時間の経過に

伴う前記ベクトルの変化からユーザーの感情バランス状態を評価するよう構成されている、請求項1記載のシステム。

【請求項3】前記認知歪み検出モジュールは、ユーザー発話内のフレーズを分析して「白黒思考」、「破局化」、「過度の自己批判」等の予め定義された認知の歪みカテゴリーに分類し、該分類結果を内部フラグとして前記対話生成モジュールに提供するように構成されている、請求項1記載のシステム。

【請求項4】前記対話生成モジュールは、認知行動療法(CBT)の手法に基づく対話スクリプトに従って応答を生成するよう構成されており、検出された認知の歪みに対して間接的に認知再構成を促す質問または提案を前記応答に含めるよう動作する、請求項1記載のシステム。

【請求項5】前記アバター出力モジュールは、3Dモデルのバーチャルアバターを用いてユーザーに視覚的応答を提供し、さらにテキスト応答を音声合成した後リアルタイム音声変換して感情トーンを付与した音声をユーザーに出力するよう構成されている、請求項1記載のシステム。

【請求項6】前記ウェアラブルセンサはスマートウォッチまたはヘルスバンドであり、心拍数、血圧、皮膚電位、睡眠パターンの少なくとも一つを測定し、該測定データがリアルタイムに前記多次元感情分析モジュールに取り込まれてユーザーの精神状態評価に反映される、請求項1記載のシステム。

【請求項7】前記分散処理手段では、ユーザーデバイス側で音声認識、表情解析およびアバター表示を実行し、クラウドまたはサーバ側で自然言語処理による感情分析および応答生成を実行するよう役割分担されている、請求項1記載のシステム。

【請求項8】前記セキュリティモジュールは、システム内のAIモデルおよびユーザーデータに対し暗号化および電子署名による保護を施し、認証されない改変が検出された場合に当該システムの動作を停止または限定モードに切り替える手段を含む、請求項1記載のシステム。

【請求項9】システムの対話管理に倫理的制約ルールを設け、宗教的・政治的勧誘や心理的な操作に該当するコンテンツを出力しないようフィルタリングするコンポーネントをさらに備える、請求項1記載のシステム。

【請求項10】 個々のユーザーの履歴データおよび嗜好情報を学習し、対話の内容や表現をユーザーごとにパーソナライズするパーソナライゼーションモジュールをさらに備える、請求項1記載のシステム。