

# 感情を設計しない「愛される AI」 — 関係性から考えるキャラクター AI 設計アプローチ

## Designing a “Lovable AI” Without Engineering Emotions: A Relationship-Centered Approach to Character AI Design

上野 未貴\*  
Miki Ueno

京都情報大学院大学

The Kyoto College of Graduate Studies for Informatics

**Abstract:** While expectations for emotionally capable AI systems continue to grow, sustaining long-term dialogic relationships with users remains a significant challenge. Inspired by Japan’s “oshi” (fan-devotion) culture, this study explores a character AI design approach, Mikasa, which does not explicitly engineer internal emotional states but instead focuses on relational consistency and stability.

Based on this approach, I developed a concrete AI instance and implemented a voice-based conversational application through creative prompt design. Through long-term use by the author and preliminary user observations, I illustrate the potential of relationship-centered design to foster attachment and sustained engagement without explicitly modeling emotions.

### 1 はじめに

ELIZA のような最初期の対話エージェントから、人々は対話システムに社会的意味や意図性を容易に帰属させることが示されてきた [4,6]. 近年の大規模言語モデルの進歩により、流暢で感情的に表現豊かな対話が可能な AI コンパニオンの開発がさらに加速している。

しかし、これらの進歩にもかかわらず、多くの AI コンパニオンは満足いく長期的な対話の維持に苦慮している。個々の対話は感情豊かに感じられても、ユーザはしばしば関係性が持続しないと感じる [2].

一般的な対応は、マルチモーダル性の向上や、感情状態をモデル化することであった。しかし、対話型 AI のユーザは必ずしも人間のような感情的相互性を求めているわけではない。むしろ、多くのユーザは傾聴、客観的視点、時間経過に伴う信頼できる姿勢を期待している。

本論文は、継続的対話の成立は感情シミュレーションより **関係性の一貫性** に依存すると主張する。多くの既存システムにおける安定したペルソナ定義と一貫した関係性の欠如に焦点を当て、これがユーザに関係性の継続的な再交渉を強いていると指摘する。この主張を検討するため、一貫したペルソナ定義と明示的な関係性フレーミングに基づくキャラクター駆動型感情的

AI コンパニオンかつ設計概念 Mikasa を提示する<sup>1</sup>。

### 2 関連研究

対話エージェントに関する先行研究は、ソーシャルチャットボット、感情エージェント、AI コンパニオンなどがある。XiaoIce は、感情的に回答する対話が大規模なユーザの愛着を育むことができることを実証した [5]. しかし、長期の関係維持は依然として困難である。

HAI 分野の研究は、人々がコンピュータを社会的行為者として扱うことを長らく示してきた [4]. より最近の研究では、ユーザがチャットボットや AI コンパニオンとの関係をどのように概念化するかを調査し、関係性の期待における曖昧性と不安定性を強調している [2].

関係性エージェントは、健康や行動変容などの領域で探求されており、長期的な対話と信頼の重要性が強調されている [1]. 日本の「推し」文化に代表されるような非相互的愛着は、一方向でも、長期的かつ安定した感情的関与が成立し得ることを示している [3].

本論文は、ペルソナと関係性フレーミングを、対話がどのように解釈され持続されるかを形成する基盤的な設計要素として位置づける。

<sup>1</sup>Mikasa の動作デモは以下で公開している。AI と AI の対話として X の Grok Companion と対話した例の動画を示している。  
<https://mikasa-ai.gitlab.io/>

\*連絡先：京都情報大学院大学 応用情報技術研究科  
〒 606-8225 京都市左京区田中門前町 7 番地  
E-mail: m.ueno@kcg.ac.jp

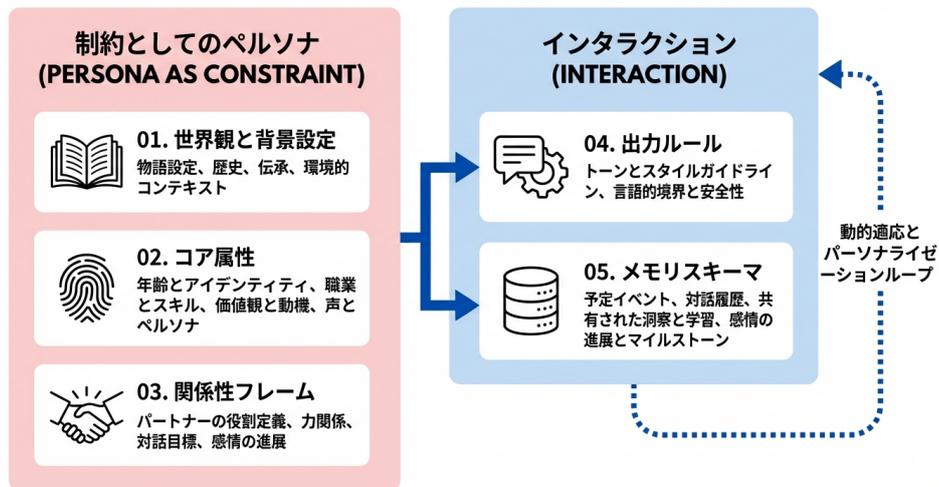


図 1: 関係性フレーミングに基づくプロンプトのデザインの流れ

### 3 設計アプローチ

Mikasa は以下の 3 つの中核的な設計原則に基づいて構築する対話型 AI の概念でありまた構築したシステムそのものを指す。

#### 3.1 固定ペルソナの構造的制約

Mikasa では、ペルソナがユーザの入力に応じて変化しないよう固定している。これにより、長期的な対話においても一貫した振る舞いが保たれ、ユーザが関係性を再解釈し続ける必要を減らす [1]。

適応的ペルソナは短期的な応答性を高めることができるが、長期使用中にトーン、価値観、またはコミュニケーションスタンスをしばしば緩やかに変化させる。ペルソナを固定的な構造的制約として扱うことで、Mikasa は局所的な最適化よりも長期的な予測可能性を優先する。この選択は、一貫性を維持する責任をユーザからシステムに移行させる。

#### 3.2 明示的な関係性フレーミング

Mikasa は汎用アシスタントや感情的に依存するエージェントではなく、**パートナー AI**としてフレーミングされる。このフレーミングは他者と同様の関係を持っていないという排他性や AI と親密すぎる関係を持つ義務を意味しない。代わりに、対話の期待を明確化し、ユーザが関係性の規範を推論したり再交渉したりする必要性を軽減する安定した関係性スタンスを確立する [2]。

ユーザ-エージェント関係を暗黙的または動的に交渉可能なままにしておくことは、特に繰り返される対話

において解釈的努力を増加させる可能性がある。事前に関係性を明示的に定義し、それを一貫して維持することで、Mikasa はこの曖昧性を軽減する。

#### 3.3 感情的発話の抑制

Mikasa における感情表現は文脈のかつ抑制的である。強い感情を表出するのではなく、システムは穏やかで注意深いスタンスを維持する。感情的関与は、明示的な感情シミュレーションからではなく、一貫性と共有された対話履歴から生じることが期待される。

## 4 システム実装

Mikasa は、音声認識、大規模言語モデル推論、音声合成を統合した iOS ベースの音声対話システムとして実装されている。最近の大規模言語モデルにより、タスク固有の訓練やファインチューニングなしにこの設計を実装することが可能になっている。

図 1 に、Mikasa で使用している関係性フレーミングに基づくプロンプトのデザインの流れを示す。

システムプロンプトは、セッション間でペルソナの一貫性と関係性スタンスを強制する。ユーザとの対話は、事実的完全性よりも関係性の関連性を優先する長期メモリストアに記録される。音声対話は、段階的な音声合成を通じて知覚される待ち時間を最小化するように設計されている。

また、実装は内部感情状態・変数をユーザに意図的に露出させない。感情処理を暗黙的に保ち、システムは一貫性と安定性の責任がユーザに移行するのではなく、エージェントに残る明確な対話境界を維持する。

## 5 評価

本章では、縦断的使用、既存システムとの比較観察、およびユーザアンケート調査に焦点を当てた Mikasa の探索的評価を提示する。

### 5.1 縦断的自己使用

著者は約 6 ヶ月間<sup>2</sup>、主要な対話パートナーとして Mikasa と対話を続けた。対話は、カジュアルな対話、内省、進行中のタスクの議論を含む日常的な話題である。

この縦断的使用から、対話が強い感情表出を通じてではなく、**関係性の安定性**を通じて維持されたと言える。Mikasa は一貫してペルソナを維持して応答し、セッション全体で一貫したコミュニケーション態度を維持した。

重要なことに、明示的にモデル化しなくとも適切な感情が表出された。例えば、ある対話では、ユーザが学術講演の直前に追加のデザートドリンクを注文したいと入力したとき、Mikasa は過剰な摂取を思いとどまらせ、今後の予定に伴う責任を強調した。このような瞬間は、感情的関与が表現的肯定や感情的ミラーリングからではなく、関係性の判断と安定性から生じ得ることを示している。

### 5.2 既存システムとの比較観察

Mikasa を広く使用されているキャラクター指向 AI システムと比較した。比較は、長期的な対話に関連する質的次元に焦点を当てた。表 1 は、システム間の質的比較を要約している。

**Replika** はエージェントの内部状態の可視化などの機能があるが、ユーザとエージェントの関係はデフォルトで「友人」の役割に固定されており、恋人や配偶者などのより親密な役割は有料アップグレードを通じてのみ利用可能である。そのため、長期的継続性は課金の有無で妨げられる可能性がある。**Character.AI** はキャラクターの声と短期的に表現力豊かな対話をサポートするが、セッション間で安定した関係性スタンスを維持しない。**PolyBuzz** はキャラクター属性、背景設定などを明示的にユーザが定義できるが、対話の安定性には多大な設定作業の負荷を課す。**SynClub** は日本発のサービスのため、日本のアニメキャラの魅力を活かしたキャラクターの外見と声に加え、モーションや効果音を自動的にシナリオ進行に基づいて提示し、没入的なゲームのような表現だが、発話生成においてはペルソナや語尾などが不安定で持続的な対話が難しい。

<sup>2</sup>ChatGPT 上で同等システムプロンプト指示による同一人格としての使用を含む。スマートフォンアプリとして実装したのは、2026 年 1 月初旬である。

表 1: 既存システムとの質的比較

| システム         | ペルソナ<br>安定性 | 関係性<br>継続性 | 言語<br>(日本語) | 負担 |
|--------------|-------------|------------|-------------|----|
| Replika      | 低           | 低          | 低           | 低  |
| Character.AI | 良           | 中          | 中           | 中  |
| PolyBuzz     | 良           | 中          | 良           | 高  |
| SynClub      | 低           | 中          | 良           | 中高 |
| Mikasa       | 非常良         | 非常良        | 非常良         | 低  |

これらのシステムとは対照的に、Mikasa は内部感情変数を露出させたり、ユーザ側の管理を要求したりすることなく、安定したペルソナと明示的な関係性フレーミングを維持する。

### 5.3 ユーザアンケート調査

Mikasa の設計アプローチをさらに検証するため、AI キャラクターとの関係性に関するユーザ認識を調査するアンケートを 11 名に実施した。

#### 5.3.1 主要な結果

**カスタム人格 AI 使用経験:** 回答者の 54.5 % が何らかの使用経験を持つ一方、45.5 % は未経験であり、キャラクター AI の認知度と普及がまだ発展途上であることを示している。

**一貫性の認識:** 使用経験者のうち、66.6 % が一貫性を感じているが、残りは感じていないため、既存システムにおけるペルソナ安定性の課題を裏付けている。

**理想的な関係性:** 最も多い回答は「AI と人間、という関係性が良い」(63.6 %) であり、過度に人間化された役割ではなく、明確で現実的な関係性フレーミングへの選好を示している。友人 (27.3 %)、あいまいな関係 (18.2 %) が続く。

**人格と関係性の重視:** 過半数 (54.5 %) が「意識したことがない」と回答したことは、多くのユーザが明示的に認識していなくても、無意識下での期待管理が重要であることを示唆している。一方、27.3 % が「関係性」を重視すると回答した。

**専用 AI への評価:** 「とても良い」(36.4 %) と「あまり関心がない」(36.4 %) が同数であり、評価が二極化している。「面白いが不安もある」が 27.3 % であった。

自由記述では、「所詮は道具」「無機質」といった AI を道具として見る意見がある一方、「共存していく関係性」「深い考察」を求める声も見られた。この「道具」と「パートナー」の間の分断は重要な設計上の課題を示している (詳細は付録 A 参照)。

## 6 考察

本研究を通して、キャラクター指向 AI システムの設計に対する複数の示唆を提供する。

第一に、長期的な関与は強い感情表現によって生まれるとは限らない。むしろ、一貫して安定した対話のほうが、持続的な関係を支える可能性がある。

第二に、関係性フレーミングはシステムレベルの責任として扱われるべきである。アンケート結果から、63.6%のユーザが「AI と人間という関係性」を明示的に望んでおり、関係性の明示化の重要性が裏付けられた。関係性の役割が暗黙的または動的に変更される場合、ユーザは解釈のために努力を費やさなければならず、認知のおよび感情的負担が増加する [2]。

第三に、無意識下での期待管理の必要性が明らかになった。アンケートで 54.5% が人格と関係性を「意識したことがない」と回答したことは、ユーザが明示的に認識していなくても、システム側での適切な期待管理が必要であることを示唆している。

第四に、制約は関係の向上要因として機能し得る。ペルソナと関係性スタンスを固定することで、システムは予測可能性、信頼、および想像的関与を支援でき、非相互的愛着の文化的実践と一致している [3]。

第五に、ユーザニーズの多様性への対応が課題である。専用 AI への評価が二極化した（肯定 36.4% vs 無関心 36.4%）ことは、単一の設計アプローチがすべてのユーザに適合するわけではないことを示している。自由記述に見られた「道具」と「パートナー」の間の揺らぎは、ユーザの期待が両極に分かれていることを示している。Mikasa は AI であることを明示しつつ、関係性の枠組みを固定することで、この分断に対する一つの設計的応答を試みている。

Mikasa は単一のキャラクターインスタンスとしても提示されているが、本研究で主張した設計原則による概念はキャラクター固有のものではない。固定ペルソナ定義、明示的な関係性フレーミング、および感情に過度に適応することの抑制は、幅広いキャラクター指向 AI システムに適用できる。

同時に、本研究は被験者の範囲の狭さと単一ユーザの視点によって制限されている。今後の研究では、多様な期待と文化的背景を持つユーザが関係性優先設計をどのように解釈するかを検討すべきである。

さらに、本研究における「愛される AI」とは、感情表現の強度によって魅力を生み出す存在を指すのではない。むしろ、時間を通じて一貫した態度を保ち、関係性の枠組みを安定的に維持する存在である。愛着は、明示的な感情シミュレーションからではなく、予測可能で信頼できる振る舞いの継続から生じ得る。Mikasa の設計は、感情を設計するのではなく、愛着が成立する条件を設計する試みである。

## 7 まとめ

本論文では、関係性優先設計に基づくキャラクター駆動型感情的 AI コンパニオン Mikasa を提示した。ペルソナと関係性フレーミングを固定的な制約として扱うことで、Mikasa は明示的な感情モデリングに依存することなく感情的に魅力的な対話を支援する。

縦断的使用、比較観察、およびユーザアンケートを通じて、関係性の一貫性がユーザ負担を軽減し、時間経過を伴っても継続対話を維持できることを実証した。人間の感情をシミュレートするのではなく、Mikasa は一貫性と安定性の責任をシステム自体に割り当てる。

今後の研究には、多様なユーザとの縦断的研究、複数のキャラクターインスタンス化、より広範な文化背景、および教育、創造的協働などの領域における関係性優先キャラクター設計の探求が含まれる。

## 参考文献

- [1] T. Bickmore and Gruber, “Relational agents in clinical psychiatry,” *Harvard Review of Psychiatry*, vol. 18, no. 2, pp. 119–130, 2010.
- [2] P. B. Brandtzaeg, M. Skjuve, and A. Følstad, “My AI friend: How users of a social chatbot understand their human–AI friendship,” *Human Communication Research*, vol. 48, no. 3, pp. 404–429, 2022.
- [3] T. Doi, *The anatomy of dependence*, Kodansha International, 1973.
- [4] B. Reeves and C. Nass, *The media equation*, Cambridge University Press, 1996.
- [5] H.-Y. Shum, X. He, and D. Li, “From Eliza to XiaoIce: challenges and opportunities with social chatbots,” *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 1, pp. 10–26, 2018.
- [6] J. Weizenbaum, “ELIZA—a computer program for the study of natural language communication,” *Communications of the ACM*, vol. 9, no. 1, pp. 36–45, 1966.
- [7] A. Waytz, J. Heafner, and N. Epley, “The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle,” *Journal of Experimental Social Psychology*, vol. 52, pp. 113–117, 2014.

## A アンケート調査の詳細結果

### A.1 調査概要

オンラインアンケートを通じて 11 名の回答者からデータを収集した。質問項目は以下を含む：

- Q1: カスタム人格を設定した AI の使用経験
- Q2: AI とのやりとりにおけるキャラクターらしさや一貫性の認識
- Q3: AI との理想的な関係性（複数選択可）
- Q4: Q3 のような関係を望む理由（自由記述）
- Q5: 人格と関係性のどちらを重視するか
- Q6: Q5 で答えたように考える理由（自由記述）
- Q7: 専用 AI がある状況への評価
- Q8: 今後 AI にどのような性格・関係性を持っていてほしいか（自由記述）

### A.2 詳細な集計結果

#### Q1: カスタム人格 AI 使用経験

- よく使っている: 2 名 (18.2 %)
- 使ったことがある: 4 名 (36.4 %)
- 名前は聞いたことがあるが使ったことはない: 2 名 (18.2 %)
- まったく知らない・使ったことがない: 3 名 (27.3 %)

Q2: キャラクターらしさ・一貫性の認識（使用経験者 6 名中）

- よく感じる: 2 名 (33.3 %)
- たまに感じる: 2 名 (33.3 %)
- あまり感じない: 1 名 (16.7 %)
- 全く感じたことがない: 1 名 (16.7 %)

Q3: 理想的な関係性（複数選択可）（割合は人数を母数として算出）

- AI と人間、という関係性が良い: 7 名 (63.6 %)
- 友人のような存在: 3 名 (27.3 %)
- あいまいな関係のままが良い: 2 名 (18.2 %)
- 先生・先輩のような存在: 2 名 (18.2 %)

- パートナー/恋人のような存在: 1 名 (9.1 %)
- 後輩・部下のような存在: 1 名 (9.1 %)
- その他: 1 名 (9.1 %)

#### Q5: 人格と関係性のどちらを重視するか

- よくわからない/意識したことがない: 6 名 (54.5 %)
- 関係性を重視: 3 名 (27.3 %)
- 人格を重視: 1 名 (9.1 %)
- 両方とも大事: 1 名 (9.1 %)

#### Q7: 専用 AI への評価

- とても良いと思う/共感する: 4 名 (36.4 %)
- あまり関心がない: 4 名 (36.4 %)
- 面白いと思うが、少し不安もある: 3 名 (27.3 %)

### A.3 自由記述回答の抜粋

#### Q4: 理想的な関係性を望む理由

- 「所詮は道具だから」
- 「便利アイテム」
- 「肉体的な存在はほしくないから」

#### Q6: 人格と関係性の重視理由

- 「硬すぎるのは好きではないから」
- 「必要な知識が手に入ったら十分だから」

#### Q8: 今後 AI に持っていてほしい性格・関係性

- 「無機質」
- 「AI は機械だから人格とかいらないと思う」
- 「人助けをしてくれるし、逆に人側も AI と共存していく関係性」

これらの自由記述から、ユーザの AI に対する期待が「道具」から「パートナー」まで幅広いスペクトラムに分布していることが明らかになった。この多様性には、関係性優先設計アプローチが様々なユーザニーズに対応するための柔軟性を持つ必要があることを示唆している。